

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 69.5142 Seconds
(without alignments)
3262.936 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534

Sequence: 1 MAGGAWGRILACYLEFLKKEE.....HLIMELWEKSKKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7534	100.0	1429	23	Human CARD-7 polyp
2	7534	100.0	1429	23	Human leucine rich
3	7534	100.0	1429	23	Leucine-rich repea
4	7534	100.0	1429	23	Human leucine-rich
5	7534	100.0	1429	24	Human caspase recr
6	7534	100.0	1429	24	Human caspase recr
7	7502	99.6	1473	22	Human NB-ARC and C
8	7502	99.6	1473	22	Human NAC beta iso
9	7488	99.4	1429	23	Human caspase recr

10	7488	99.4	1429	23	AAO17855	Pyrin domain conta
11	7484	99.3	1473	22	AAE06758	Human G-protein co
12	7342	97.5	1397	22	RAY72670	Human NB-ARC and C
13	7325.5	97.2	1442	22	AAAY72671	Human NB-ARC and C
14	6344	84.2	1454	22	AAAY72673	Human NAC beta iso
15	6174	81.9	1424	22	AAAY72674	Human NAC gamma or
16	2171.5	28.8	442	21	AAAB24513	Human secreted pro
17	1380	18.3	1061	23	AAO15590	Human PYRIN-8 prot
18	1375	18.3	1099	23	AAO17857	Pyrin domain conta
19	1279	17.0	1034	22	AAE07514	Human PYRIN-1 prot
20	1279	17.0	1034	24	ABU08503	Huma PYRIN-1 prote
21	1263.5	16.8	920	23	ABP53254	Human MDDT-13 prot
22	1154	15.3	896	23	AAO15592	Human PYRIN-11 prot
23	1037	13.8	994	23	AAO15593	Human PYRIN-3 prot
24	1020.5	13.5	983	23	AAO17870	Pyrin domain conta
25	1007	13.4	919	23	ABP43483	Human secreted pro
26	1007	13.4	919	24	ABJ19319	NOVX related prote
27	985.5	13.1	1162	23	AAE25053	Human EMERY-1 prot
28	983	13.0	1200	23	AAU79526	Human EMERY-1 prot
29	983	13.0	1200	24	AAE31749	Human MATER protei
30	982.5	13.0	2312	23	AAO17859	Pyrin domain conta
31	972.5	12.9	858	23	AAO17859	Human nucleotide b
32	948	12.6	1062	22	AAAG65895	Amino acid sequenc
33	948	12.6	1062	23	AAO17856	Pyrin domain conta
34	945.5	12.5	1033	22	AAE07513	Human nucleotide b
35	945.5	12.5	1033	23	ABG78456	Human nucleotide b
36	945.5	12.5	1033	24	ABU08502	Huma pyrin domain
37	945.5	12.5	1344	23	AAO15585	Human PYRIN-5 prot
38	941.5	12.5	1049	22	AAAG65894	Amino acid sequenc
39	934.5	12.4	1033	23	ABU65214	Human NOV125b prot
40	924.5	12.3	732	23	AAO17858	Pyrin domain conta
41	922.5	12.2	952	22	ABG03924	Novel human diagno
42	910	12.1	1022	23	AAO17862	Pyrin domain conta
43	908.5	12.1	980	23	ABB98244	Human CGDP SQ ID
44	908.5	12.1	980	23	AAAM50329	Human nucleotide b
45	897	11.9	891	22	AAE04546	Human G-protein co

ALIGNMENTS

RESULT 1
AAB62571
ID AAB62571 standard; Protein; 1429 AA.

XX AC AAB62571;

XX 23-JUL-2001 (first entry)

DT Human CARD-7 polypeptide.

XX CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;

DE autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;

KW antibacterial; antiviral; gene therapy.

XX Homo sapiens.

XX Key

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

Location/Qualifiers
329..645
/note="nucleotide binding domain"
333..341
/note="kinase 1A subdomain"
404..413
/note="kinase 2 subdomain"
454..463
/note="kinase 3a subdomain"
615..622
/note="motif 2 domain"
807..834
/note="leucine-rich domain"
836..863
/note="leucine-rich domain"
864..891

FT Domain /note= "leucine-rich domain"
 FT 893..920
 FT /note= "leucine-rich domain"
 FT 921..948
 FT /note= "leucine-rich domain"
 FT 950..976
 FT /note= "leucine-rich domain"
 XX WO200130813-A1.
 PN 03-MAY-2001.
 XX 27-OCT-2000; 2000WO-US29796.
 XX 27-OCT-1999; 99US-0428252.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Bertin J;
 XX WPI: 2001-343340/36.
 DR N-P8DB; AAF83651.
 XX
 XX Identifying a modulator of interaction between caspase recruitment
 PT domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
 PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of
 PT the compound -
 XX
 PS Disclosure; Fig 1A-D; 80pp; English.
 XX
 CC The invention relates to identifying compounds, that modulate interaction
 CC between caspase recruitment domain (CARD)-7 and CARD-5. The method
 CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
 CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
 CC presence of the compound compared to the binding in the absence of the
 CC compound indicates that the compound is a modulator of CARD-7-CARD-5
 CC interaction). Modulators of CARD-7 and CARD-8 expression or activity can
 CC be used to treat or diagnose disorders such as cancers, bacterial or
 CC viral infections, autoimmune disorders (systemic lupus erythematosus,
 CC immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
 CC organ-specific autoimmunity including multiple sclerosis, Hashimoto's
 CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
 CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
 CC of cellular processes including cell growth and cell death. The present
 CC sequence represents the human CARD-7, an intracellular protein.
 XX
 SQ Sequence 1429 AA;
 Query Match 100.0%; Score 7534; DB 22; Length 1429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSGMVEASVLAQ 60
 DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSGMVEASVLAQ 60
 QY 61 YGEQAWDLALHTWQMGRLSICAQAGAGHSPSPSPHGLSPSQPTSTAVLMPW 120
 DB 61 YGEQAWDLALHTWQMGRLSICAQAGAGHSPSPSPHGLSPSQPTSTAVLMPW 120
 QY 121 IHELPACTQGSERRVLRQLPTSGRRWREISASLLYQALPSPDPHESPSQSPNAPTST 180
 DB 121 IHELPACTQGSERRVLRQLPTSGRRWREISASLLYQALPSPDPHESPSQSPNAPTST 180
 QY 181 AVLGSWGPQPSLAPREQEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
 DB 181 AVLGSWGPQPSLAPREQEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
 QY 241 PQAHTSLOPHHPHWPSPRESLCSWPKNEDFNQKFTQLLLLOPHRPSODPLVKRSW 300
 DB 241 PQAHTSLOPHHPHWPSPRESLCSWPKNEDFNQKFTQLLLLOPHRPSODPLVKRSW 300
 QY 301 PDYVEENRGHLIEIRDLFGPGLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGR 360

DB PDYVEENRGHLIEIRDLFGPGLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGR 360
 QY FQHVYFSCRELAQSKVVSIAELIGKDGATAPAPIRQILSRPERLLFLDGDVDEPGWLQ 420
 DB FQHVYFSCRELAQSKVVSIAELIGKDGATAPAPIRQILSRPERLLFLDGDVDEPGWLQ 420
 QY EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSLQARVVEVLG 480
 DB EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSLQARVVEVLG 480
 QY FSESRKEYFYRYFTDERQAIAPRLVKSNNKELNALCLVPVWVWMLACTCLQOMKRKEKL 540
 DB FSESRKEYFYRYFTDERQAIAPRLVKSNNKELNALCLVPVWVWMLACTCLQOMKRKEKL 540
 QY TLTSKTTTTLCHVLAQALQAPLGPQRLDLSLAAGIWOQKTLFSPDDLRKHGLDCAI 600
 DB TLTSKTTTTLCHVLAQALQAPLGPQRLDLSLAAGIWOQKTLFSPDDLRKHGLDCAI 600
 QY ISTFLKMGILQEHPIPLSYFIHLFCQEFFAAMSIVLEDEKGRGHSHNCIIDLEKTL 660
 DB ISTFLKMGILQEHPIPLSYFIHLFCQEFFAAMSIVLEDEKGRGHSHNCIIDLEKTL 660
 QY GIHGLFGASTTRFLGLLSDDEGEREMENIFHCRLSQGRNLMQWVPSLQQLLOPHSLSLH 720
 DB GIHGLFGASTTRFLGLLSDDEGEREMENIFHCRLSQGRNLMQWVPSLQQLLOPHSLSLH 720
 QY CLYETRNKTFITQWMAHFEENGMCVETDMLLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780
 DB CLYETRNKTFITQWMAHFEENGMCVETDMLLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780
 QY PTMVLFRWVPVTDAYWOLFSLVTKNKLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
 DB PTMVLFRWVPVTDAYWOLFSLVTKNKLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
 QY TRLAGCGLTAEDCKDLAFGLRANQTLTDLNLSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
 DB TRLAGCGLTAEDCKDLAFGLRANQTLTDLNLSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
 QY LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLRLGLDQ 960
 DB LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDVGVRLLCGLRHPACKLRLGLDQ 960
 QY TTLSDEMRQELRALOEKPOLLIIFSRKPSVMTTPEGLDTCGEMNSSTSLKRLGSR 1020
 DB TTLSDEMRQELRALOEKPOLLIIFSRKPSVMTTPEGLDTCGEMNSSTSLKRLGSR 1020
 QY ASHVAQANLKLDDVSKIPPIAIEAESSEPEVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080
 DB ASHVAQANLKLDDVSKIPPIAIEAESSEPEVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080
 QY GTPGVATEVVDKKNLYRVHFPVAGSYRWNTGICFVWREAVTVEIFCVMQDQFLGIN 1140
 DB GTPGVATEVVDKKNLYRVHFPVAGSYRWNTGICFVWREAVTVEIFCVMQDQFLGIN 1140
 QY PQSHMWAGPLDLDKABEGAVEAHLPHFVALQGHVDTSLFQMAHFKEGMLEKPARV 1200
 DB PQSHMWAGPLDLDKABEGAVEAHLPHFVALQGHVDTSLFQMAHFKEGMLEKPARV 1200
 QY ELHIVLENPSFPLGVLKMIHNAIRFIPVTSVVLLYHRVHPPEVTPHLYLIPSDCSIR 1260
 DB ELHIVLENPSFPLGVLKMIHNAIRFIPVTSVVLLYHRVHPPEVTPHLYLIPSDCSIR 1260
 QY KELELCYRSPQEDQFSEFYVGHGSGIRLQVKDKDETLVWEALVKPGDLMPTLIPP 1320
 DB KELELCYRSPQEDQFSEFYVGHGSGIRLQVKDKDETLVWEALVKPGDLMPTLIPP 1320
 QY ARIAVPSPLDAPQLLHFVDQVREQLIARVTSVEVLDKLGQVLSQEOYERVLANTRPS 1380
 DB ARIAVPSPLDAPQLLHFVDQVREQLIARVTSVEVLDKLGQVLSQEOYERVLANTRPS 1380
 QY QMRKLFSLISQSWDRCKDKGLYQALKETHPHILIMELWEKSGSKGLPLSS 1429

Db	1381 QMRKLFSLSQSDRKCKDGLYOALKETHPHLIMELWEKSGKGLPLSS 1429	Db	1 MAGGAWGRACYLEFLKKEELKEFOLLANKAHSSSSGETPAQPEKTSGMEVASYLVAQ 60
RESULT 2		QY	61 YGEORANDLALHTWEQMGRLSLCAQAQEGAGHSFPYPSPSEPHLGSPSQSTSTAVLMPW 120
ABG97969		Db	61 YGEORANDLALHTWEQMGRLSLCAQAQEGAGHSFPYPSPSEPHLGSPSQSTSTAVLMPW 120
AC	ABG97969 standard; Protein; 1429 AA.	QY	121 IHELPAQCTQSERRVLRLPDTSGRRWRREISASLLYQALPSSPDHSPDESNPAPTST 180
XX	ABG97969;	Db	121 IHELPAQCTQSERRVLRLPDTSGRRWRREISASLLYQALPSSPDHSPDESNPAPTST 180
DT	07-JAN-2003 (first entry)	QY	181 AVLSGWSGPPPSLAPREOEAPGTOWPLDETSGIYVTEIREREREKSKGKPPWAAVGT 240
XX	Human leucine rich repeat domain containing protein #2.	Db	181 AVLSGWSGPPPSLAPREOEAPGTOWPLDETSGIYVTEIREREREKSKGKPPWAAVGT 240
DE		QY	241 PPQAHSTLQPHHHPEPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPRQDPLVKESW 300
KW	Leucine rich repeat; nervous system; human; neural disorder; apoptosis;	Db	241 PPQAHSTLQPHHHPEPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPRQDPLVKESW 300
KW	renal disorder; immune disorder; arthritis; asthma; AIDS;	QY	301 PDYVEENRGHLIEIRDLFGPGLDTPQEPRIVLQGAAGIKSTLARQVKEANGRGOLYGR 360
KW	acquired immunodeficiency syndrome; rheumatoid arthritis;	Db	301 PDYVEENRGHLIEIRDLFGPGLDTPQEPRIVLQGAAGIKSTLARQVKEANGRGOLYGR 360
KW	haematopoietic disorder; metabolic disorder; reproductive disorder;	QY	361 FOHVPFSCRELAQSKVVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWVLQ 420
KW	pulmonary disease; cardiovascular disease; hyperproliferative disorder;	Db	361 FOHVPFSCRELAQSKVVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWVLQ 420
KW	neurodegenerative disease; Alzheimer's disease; Parkinson's disease;	QY	421 EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSEQARVVEVLG 480
KW	Huntington's disease; developmental disorder; autoimmune disease;	Db	421 EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSEQARVVEVLG 480
KW	Addison's disease; haemolytic anaemia; antiphospholipid syndrome;	QY	481 FSESRKEYFYRYFTDERQAIAPRLVKSNEKELWALCLVPVWVSWLACTCLMQMKRKEKL 540
KW	allergic encephalomyelitis; gene therapy.	Db	481 FSESRKEYFYRYFTDERQAIAPRLVKSNEKELWALCLVPVWVSWLACTCLMQMKRKEKL 540
XX		QY	541 TLTSKTTTTLCLHYLAQALQPLGPOLRDLCSLAEGIWOKKTLFSPDDLRKHGLDGA 600
OS	Homo sapiens.	Db	541 TLTSKTTTTLCLHYLAQALQPLGPOLRDLCSLAEGIWOKKTLFSPDDLRKHGLDGA 600
XX		QY	601 ISTFLKMGILQEHPIPLSYFIHLFCQEFFAAMSIVLEDEKGRGHKSHNCIIIDLEKTL 660
PN	WO200274959-A2.	Db	601 ISTFLKMGILQEHPIPLSYFIHLFCQEFFAAMSIVLEDEKGRGHKSHNCIIIDLEKTL 660
XX	26-SEP-2002.	QY	661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLQMWVPSLQLLQPHSLES 720
XX		Db	661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLQMWVPSLQLLQPHSLES 720
PF	20-DEC-2001; 2001WO-US50457.	QY	721 CLYETRNKTLTQVMAHEEMCMCVETDMLLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780
XX		Db	721 CLYETRNKTLTQVMAHEEMCMCVETDMLLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780
PR	03-JAN-2001; 2001US-259479P.	QY	781 PTMVVLFWRVPVTDAYWQILFSLVLRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
PR	09-JAN-2001; 2001US-260616P.	Db	781 PTMVVLFWRVPVTDAYWQILFSLVLRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
XX	(BRIM) BRISTOL-MYERS SQUIBB CO.	QY	841 TLRLAGCGLTAECDKDLAFGLRANQTLFELDSFNVLTDAGAKHLQRLRQPSCKLQRIQ 900
PA		Db	841 TLRLAGCGLTAECDKDLAFGLRANQTLFELDSFNVLTDAGAKHLQRLRQPSCKLQRIQ 900
XX		QY	901 LVSCGLTSDCCODLASVLSASPSLKLDELQNNLDVGVRLICEGRHPACKLIRGLDQ 960
XX		Db	901 LVSCGLTSDCCODLASVLSASPSLKLDELQNNLDVGVRLICEGRHPACKLIRGLDQ 960
PI	Ramanathan C, Feder J, Mintier G;	QY	961 TTLSDEMQELRALRQEKPOLLI FSRKRPVMTPTTEGLDTGEMSNSTSLSKRQLGSE 1020
XX		Db	961 TTLSDEMQELRALRQEKPOLLI FSRKRPVMTPTTEGLDTGEMSNSTSLSKRQLGSE 1020
DR	WPI; 2002-750554/81.	QY	1021 ASHVAQANLKLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGLHTKPLGTD 1080
XX		Db	1021 ASHVAQANLKLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGLHTKPLGTD 1080
PT	New HLRNS1 nucleic acids and polypeptides, useful for preventing,	QY	1081 GPTGPVATEVDKKNLVRVHPFPVAGSRWPNVTGLCFVNRVATVTEIFCVDQFLGEIN 1140
PT	treating, or ameliorating e.g. renal disorder, immune, hematopoietic,	Db	1081 GPTGPVATEVDKKNLVRVHPFPVAGSRWPNVTGLCFVNRVATVTEIFCVDQFLGEIN 1140
PT	metabolic, reproductive, pulmonary, cardiovascular or autoimmune		
PT	diseases		
XX	Example 1; Page 392-397; 415pp; English.		
PS			
XX	The invention describes nucleic acids encoding human leucine-rich repeat		
CC	containing proteins expressed in nervous system tissues, HLRNS1. The		
CC	HLRNS1 polypeptide or the polynucleotide is useful for preventing,		
CC	treating, or ameliorating a neural disorder or a disorder related to		
CC	aberrant apoptosis modulation (either directly or indirectly), renal		
CC	disorder, immune disorder (e.g. arthritis, asthma, acquired		
CC	immunodeficiency syndrome (AIDS) or rheumatoid arthritis),		
CC	haematopoietic, metabolic, reproductive, pulmonary or cardiovascular		
CC	diseases, hyperproliferative disorders, neurodegenerative diseases		
CC	(e.g. Alzheimer's disease, Parkinson's disease or Huntington's		
CC	disease), developmental disorders, non-infections disorders, nervous		
CC	system diseases and/or disorders, and autoimmune diseases (e.g.		
CC	Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or		
CC	allergic encephalomyelitis). The polynucleotides are also useful as		
CC	chromosome markers, for chromosome identification, gene therapy, and in		
CC	identifying organisms from minute biological samples. This is the amino		
CC	acid sequence of a leucine-rich repeat containing protein.		
XX			
SQ	Sequence 1429 AA;		
	Query Match 100.0%; Score 7534; DB 23; Length 1429;		
	Best Local Similarity 100.0%; Pred No. 0;		
	Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MAGGAWGRACYLEFLKKEELKEFOLLANKAHSSSSGETPAQPEKTSGMEVASYLVAQ 60		

QY 1141 POHSMWVAGPLLDIKAEPAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
DB 1141 POHSMWVAGPLLDIKAEPAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
QY 1201 ELHHIVLENPSPPLGVLLKMIHNAURFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
DB 1201 ELHHIVLENPSPPLGVLLKMIHNAURFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
QY 1261 KELELCYRSPGEDOLFSEFVYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
DB 1261 KELELCYRSPGEDOLFSEFVYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
QY 1321 ARTAVPSPLDAPOLLHFVQYREOLIAVTSVEVLDLKLHGOVLSQEQYERVLAEINTPS 1380
DB 1321 ARTAVPSPLDAPOLLHFVQYREOLIAVTSVEVLDLKLHGOVLSQEQYERVLAEINTPS 1380
QY 1381 QMRKLFSLQSQSDRKCKDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
DB 1381 QMRKLFSLQSQSDRKCKDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
RESULT 3
ID ABG78472
XX ABG78472 standard; Protein; 1429 AA.
AC ABG78472;
XX
DT 15-NOV-2002 (first entry)
DE Leucine-rich repeat (LRR) domain containing protein #2.
KW Human; human leucine-rich repeat small intestine I; HLRRS11; asthma;
KW proliferative disorder; gastrointestinal disorder; renal disorder;
KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
KW blood coagulation disorder; autoimmune disorder; allergic reaction;
KW inflammatory condition; cardiovascular disorder; ischaemia;
KW neurological disorder; infectious disease; cytokine production;
KW expressed sequence tag; EST.
XX
OS Unidentified.
PN WO200261086-A2.
XX
PD 08-AUG-2002.
PF 20-DEC-2001; 2001WO-US49739.
XX
PR 22-DEC-2000; 2000US-257774P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder J, Ramanathan C, Mintier G;
XX WPI; 2002-619252/66.
XX
PT New isolated nucleic acid molecules encoding HLRRS11 polypeptides, or
PT their fragments and homologues, useful for preventing, treating and
PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
PT or renal disorders
XX
PS Disclosure; Page 229-233; 336pp; English.
XX
CC The invention relates to isolated nucleic acid molecules (I) encoding
CC human leucine-rich repeat small intestine I (HLRRS11) polypeptides.
CC The nucleic acid molecules and polypeptides are useful for preventing,
CC treating and ameliorating medical conditions, such as proliferative,
CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
CC related to aberrant calcium regulation or apoptosis modulation, either
CC directly or indirectly. They are also useful for treating, preventing
CC and/or diagnosing diseases, disorders and/or conditions of: immune system
CC by activating or inhibiting the proliferation, differentiation, or

CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
CC virus (HIV) infection, HTLV-BIV infection; blood coagulation disorders,
CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
CC myasthenia gravis; asthma or allergic reactions; inflammatory
CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
CC pneumonia, or viral, bacterial, and fungal infections. The HLRRS11
CC polypeptides are useful for modulating cytokine production, antigen
CC presentation, or other processes such as boosting immune responses.
CC ABG78454-ABG78474 represent HLRRS11 amino acid sequences and related
CC amino acid sequences of the invention.
XX
SQ Sequence 1429 AA;
Query Match 100.0%; Score 7534; DB 23; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSKGHEVASLYVAQ 60
DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSKGHEVASLYVAQ 60
QY 61 YGEQAWDLALHTWEOMGLRSLCAQAQEGAGHSFPYSPSEPHLGSPSQSTSTAVLMPW 120
DB 61 YGEQAWDLALHTWEOMGLRSLCAQAQEGAGHSFPYSPSEPHLGSPSQSTSTAVLMPW 120
QY 121 IHELPAGTCQSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPESNAPTST 180
DB 121 IHELPAGTCQSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPESNAPTST 180
QY 181 AVLGSWGSPPOS LAPROEAPGTQWPLDETSGIYYTIREREREKSKGKPPHAAVGT 240
DB 181 AVLGSWGSPPOS LAPROEAPGTQWPLDETSGIYYTIREREREKSKGKPPHAAVGT 240
QY 241 PPOAHTSLQPHHPWPSPVRESLCSWTWPKNEDFNQKFTQLLLQRPHPRQDPLVKRSW 300
DB 241 PPOAHTSLQPHHPWPSPVRESLCSWTWPKNEDFNQKFTQLLLQRPHPRQDPLVKRSW 300
QY 301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVLQGAAGIGKSTLARQVKEAMRGOLYQDR 360
DB 301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVLQGAAGIGKSTLARQVKEAMRGOLYQDR 360
QY 361 FOHVYFSCRELAQSKVVS LAELIGKDGATATAPAIROIILSRPERLLFLDGDVDEPGWVLQ 420
DB 361 FOHVYFSCRELAQSKVVS LAELIGKDGATATAPAIROIILSRPERLLFLDGDVDEPGWVLQ 420
QY 421 EPSSELCLHWSQOPADALLGSLGKTLTPASFLITARTTALQNLIPSLQARWVEVLG 480
DB 421 EPSSELCLHWSQOPADALLGSLGKTLTPASFLITARTTALQNLIPSLQARWVEVLG 480
QY 481 FSESRKEYFYRYFTDERQAIAPRLVKSNNELWALCLVPVWVSWLACTCLQMKRKEKL 540
DB 481 FSESRKEYFYRYFTDERQAIAPRLVKSNNELWALCLVPVWVSWLACTCLQMKRKEKL 540
QY 541 TLTSKTTTTLCHLYLAQALQAPLGPQRDLCSLAAGIWKOKTLPSPDDLRKHGLDGA 600
DB 541 TLTSKTTTTLCHLYLAQALQAPLGPQRDLCSLAAGIWKOKTLPSPDDLRKHGLDGA 600
QY 601 ISTFLKMGILQHPHPIPLSYFIHLCSFQFFAAMS VVLEDEKRGKHSNCIIDLEKTLBAY 660
DB 601 ISTFLKMGILQHPHPIPLSYFIHLCSFQFFAAMS VVLEDEKRGKHSNCIIDLEKTLBAY 660
QY 661 GIHGLFGASTTRFLIGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
DB 661 GIHGLFGASTTRFLIGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
QY 721 CLYETRNKTFUTQWMAHFEEMGCMVETDMELLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780
DB 721 CLYETRNKTFUTQWMAHFEEMGCMVETDMELLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780

QY 781 PTMVLFRRVPTDAYWQILFVLKVRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
Db 781 PTMVLFRRVPTDAYWQILFVLKVRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
QY 841 TLRLAGCGLTAECDCKDLAFGLRANOTLTFLDLSFNVLTADAGAKHLCORLROPSCKLQRLQ 900
Db 841 TLRLAGCGLTAECDCKDLAFGLRANOTLTFLDLSFNVLTADAGAKHLCORLROPSCKLQRLQ 900
QY 901 LVSCGLTSCCCODLASVLSASPSLKELDLOQNNLDDVGVRLICEGLRHPACKLIRGLDQ 960
Db 901 LVSCGLTSCCCODLASVLSASPSLKELDLOQNNLDDVGVRLICEGLRHPACKLIRGLDQ 960
QY 961 TTLSDEMROELRALSOEKPOLLIFFRRKPSVMTPEGLDTGMSNSTSLKQRIGSERA 1020
Db 961 TTLSDEMROELRALSOEKPOLLIFFRRKPSVMTPEGLDTGMSNSTSLKQRIGSERA 1020
QY 1021 ASHVAQANLKLVDVSKIFPIAIAEESSEPVVVELLCVPSASOGDLHTKPLGTDDDFW 1080
Db 1021 ASHVAQANLKLVDVSKIFPIAIAEESSEPVVVELLCVPSASOGDLHTKPLGTDDDFW 1080
QY 1081 GPTGPVATEVVDKKNLYRVHPVPVAGSYRPNVTGLCFVNRREAVTVIEIFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKKNLYRVHPVPVAGSYRPNVTGLCFVNRREAVTVIEIFCVWDQFLGEIN 1140
QY 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFOMAHFKEEGLMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFOMAHFKEEGLMLLEKPARV 1200
QY 1201 ELHHIVLENPSPGLVLLKMTNARLFTPTVSULLYHRVHPEVTHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPGLVLLKMTNARLFTPTVSULLYHRVHPEVTHLYLIPSDCSIR 1260
QY 1261 KELELCYRSPGEDQLPSEFVYVGHLSGIRLOVKDKDELTVLWELVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGEDQLPSEFVYVGHLSGIRLOVKDKDELTVLWELVKPGDLMPATTLIPP 1320
QY 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
QY 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
RESULT 4
ABB77916
ID ABB77916 standard; protein; 1429 AA.
XX AC ABB77916;
XX DT 07-OCT-2002 (first entry)
XX DE Human leucine-rich repeat domain containing protein KIAA0926.
XX KW Human; leucine-rich repeat; HLRRBM1; proliferative disorder;
KW immune condition; apoptosis; signal transduction; autoimmune disease;
KW haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
KW cardiovascular disorder; neurological disease; pheromone;
KW pulmonary disease; chronic obstructive pulmonary disease;
KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
KW haematopoietic disease; platelet disorder; Bernard-Soulier
KW inflammatory disorder; systemic lupus erythematosus;
KW cardiovascular disease; cancer; KIAA0926.
XX OS Homo sapiens.
XX PN WO200252011-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-US49740.
XX

PR 22-DEC-2000; 2000US-257773P.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Feder J, Ramanathan C, Mintier G;
XX WPI; 2002-566676/60.
DR
PT New HLRRBM1 nucleic acids for preventing, treating or ameliorating e.g.
PT proliferative disorders, immune conditions, a disorder related to
PT aberrant apoptosis modulation or developmental disorders -
XX
PS Example 1; Page 362-366; 371pp; English.
XX
CC The present sequence represents a human leucine-rich repeat domain
CC containing protein, which was used as a probe to search for leucine-rich
CC repeat containing protein HLRRBM1. HLRRBM1 polypeptides and
CC polynucleotides are useful for preventing, treating or ameliorating a
CC medical condition such as a proliferative disorder, immune condition, or
CC a disorder related to aberrant apoptosis modulation, either directly or
CC indirectly, and in modulating signal transduction activity in various
CC cells, tissue and organisms. They are also useful for treating,
CC preventing, or diagnosing diseases of haematopoietic cells, autoimmune
CC disease, graft-versus-host disease, allergic conditions (e.g. asthma),
CC cardiovascular disorders, and neurological diseases, and for increasing
CC the organisms' ability to synthesize and/or release pheromones. The
CC polypeptide may also be used in treating, preventing or ameliorating
CC pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic
CC rhinitis, or bronchial hyperresponsiveness), reproductive disease,
CC haematopoietic disease, platelet disorders (e.g. Bernard-Soulier
CC syndrome), non-infectious immune disorders (e.g. innate immunity to bacterial
CC pathogens, or adaptive immune response), immune and inflammatory
CC disorders (e.g. systemic lupus erythematosus), cardiovascular diseases
CC and cancers. HLRRBM1 nucleic acids may further be used in chromosome
CC identification or mapping, as a chromosome marker, as molecular weight
CC markers, as diagnostic probes, in gene therapy, in raising anti-DNA
CC antibodies, or as antigens for eliciting immune responses.
XX
SQ Sequence 1429 AA;
Query Match 100.0%; Score 7534; DB 23; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHRSSSGETPAQPEKTSGMVEASVYLAQ 60
Db 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHRSSSGETPAQPEKTSGMVEASVYLAQ 60
QY 61 YGEQRAWDLALHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
Db 61 YGEQRAWDLALHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
QY 121 THELPAGTCQSERRVLRQLPDTSGRRWRRETSASLLYQALPSSPDHESQSPNAPTST 180
Db 121 THELPAGTCQSERRVLRQLPDTSGRRWRRETSASLLYQALPSSPDHESQSPNAPTST 180
QY 181 AVLGSWGPPOPSLAPREQEAPGTOWPLDETSGIYYTIRERERESKSGRPPMAVVGVT 240
Db 181 AVLGSWGPPOPSLAPREQEAPGTOWPLDETSGIYYTIRERERESKSGRPPMAVVGVT 240
QY 241 PPOAHTSLQPHHHWPSPVRESLCSCTWPWKNEDFNQKFTQLLLQRPHPRSODPLVKGSW 300
Db 241 PPOAHTSLQPHHHWPSPVRESLCSCTWPWKNEDFNQKFTQLLLQRPHPRSODPLVKGSW 300
QY 301 PDYVEENRGHILIEIRDLFGPGLDTPRIVILQGAAGICKSTLARQVKEAWRGOLYGD 360
Db 301 PDYVEENRGHILIEIRDLFGPGLDTPRIVILQGAAGICKSTLARQVKEAWRGOLYGD 360
QY 361 FOHVYFSCRELAQSKVSVSLAELIGDKGTATPAPTRQILSRPERLLFLDGVDEPGWLQ 420
Db 361 FOHVYFSCRELAQSKVSVSLAELIGDKGTATPAPTRQILSRPERLLFLDGVDEPGWLQ 420
QY 421 EPSSELCLHWSQPQPADALLGSLGKLTILPEASFLITARTTALQNLIPSEQARWVEVLG 480

Db 421 EFSSSELCHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSLQARWVEVLG 480
 Qy 481 FSESSKEFYFYFTDERQAIAPRLVSKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
 Db 481 FSESSKEFYFYFTDERQAIAPRLVSKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
 Qy 541 TLTSTKTTTTLCHYLAAQALQAPLQPLQDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 Db 541 TLTSTKTTTTLCHYLAAQALQAPLQPLQDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 Qy 601 ISTFLKMGILQHEPIPLSYFHLFCQEPFAAMSVYLEDKGRGHSNCIIDLKXTLEAY 660
 Db 601 ISTFLKMGILQHEPIPLSYFHLFCQEPFAAMSVYLEDKGRGHSNCIIDLKXTLEAY 660
 Qy 661 GHGLFGASTTFLGLLSDGEREMENIFHCLSGQRNLQMWVPSLQLLQPHSLESILH 720
 Db 661 GHGLFGASTTFLGLLSDGEREMENIFHCLSGQRNLQMWVPSLQLLQPHSLESILH 720
 Qy 721 CLVETRNKTFLOVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKQLQIEGRQHRSTWS 780
 Db 721 CLVETRNKTFLOVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKQLQIEGRQHRSTWS 780
 Qy 781 PTMVVLFWRVPVTDAYWQILFVSLKVTNRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
 Db 781 PTMVVLFWRVPVTDAYWQILFVSLKVTNRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
 Qy 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 Db 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 Qy 901 LVSCGLTSCCCDLASVLSASLSKELDLQNNLDVGVRLLCCEGLRHHPACKLIRGLDQ 960
 Db 901 LVSCGLTSCCCDLASVLSASLSKELDLQNNLDVGVRLLCCEGLRHHPACKLIRGLDQ 960
 Qy 961 TTLSDEMROELRALEQEKQLLIFSRKPSVMTPTGLDTGEMSNSTSLKQRQLGSE 1020
 Db 961 TTLSDEMROELRALEQEKQLLIFSRKPSVMTPTGLDTGEMSNSTSLKQRQLGSE 1020
 Qy 1021 ASHVAQANILKLDVSKIPIAIEABESSPEVVPVELLCVPSASQDGLTKPLGTDGDFW 1080
 Db 1021 ASHVAQANILKLDVSKIPIAIEABESSPEVVPVELLCVPSASQDGLTKPLGTDGDFW 1080
 Qy 1081 GTPGPVATEVDEKKNLYRHPVPVAGSYRWPNTGLCFVNRVATVEIEFCVWDQFLGEIN 1140
 Db 1081 GTPGPVATEVDEKKNLYRHPVPVAGSYRWPNTGLCFVNRVATVEIEFCVWDQFLGEIN 1140
 Qy 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKPEGMLLEKPARV 1200
 Db 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKPEGMLLEKPARV 1200
 Qy 1201 ELHHIVLENPSPGLGVLLKMHNLARFPTVTSVLLYHRVPEEVTFLYLIIPSDCSIR 1260
 Db 1201 ELHHIVLENPSPGLGVLLKMHNLARFPTVTSVLLYHRVPEEVTFLYLIIPSDCSIR 1260
 Qy 1261 KELELCYRSPGDDQLFSEFYVGHLSGIRLQVKKDETLVNEALVKPGDLMPATTLIPP 1320
 Db 1261 KELELCYRSPGDDQLFSEFYVGHLSGIRLQVKKDETLVNEALVKPGDLMPATTLIPP 1320
 Qy 1321 ARIAVPSPLDAPQLLHFVDQYEQELIARVTSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
 Db 1321 ARIAVPSPLDAPQLLHFVDQYEQELIARVTSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
 Qy 1381 QMRKLFSLQSQWDRCKCKGLYOALKETHPHLIMELWEKSKGGLPLSS 1429
 Db 1381 QMRKLFSLQSQWDRCKCKGLYOALKETHPHLIMELWEKSKGGLPLSS 1429

RESULT 5
 ABG71631
 ID ABG71631 standard; Protein; 1429 AA.
 XX
 AC ABG71631;

XX DT 09-JAN-2003 (first entry)
 XX DE Human caspase recruitment domain-7 (CARD-7).
 XX KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudolICE; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant.
 OS Homo sapiens.
 XX US2002128219-A1.
 XX PD 12-SEP-2002.
 XX PF 15-AUG-2001; 2001US-0931071.
 XX PR 27-OCT-1999; 99US-0428252.
 XX (BERT/) BERTIN J.
 XX PA (ALNE/) ALNEMRI E S.
 XX PI Bertin J, Alnemri ES;
 XX DR WPI; 2003-028968/02.
 XX DR N-PSDB; ABS55497.
 PT Assays for identifying compound that modulates the interaction of
 PT caspase recruitment domain-8 with a CARD-8 ligand or a compound that
 PT modulates activity of CARD-8 -
 XX Disclosure; Fig 1; 49pp; English.
 XX The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudolICE) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropaenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC represents human CARD-7.

961	TTLSDBMRQELRALQEQEKPOLLI	FSRRPSPVMTPT	EGDLTGEMSNSTSSLKQRQLGSERA	1020
1021	ASHVAQANLKL	LDVSKIPPIA	IEAESPEVVPVELLCVPSPASQGLDHTKPLGTD	1080
1021	ASHVAQANLKL	LDVSKIPPIA	IEAESPEVVPVELLCVPSPASQGLDHTKPLGTD	1080
1081	GTPGPVATEVDK	KNLVRVHFPVAGSYRWPN	TGLCFVMREAVTVIEIFECVWDQFLGEIN	1140
1081	GTPGPVATEVDK	KNLVRVHFPVAGSYRWPN	TGLCFVMREAVTVIEIFECVWDQFLGEIN	1140
1141	POHSWMVAGPLD	IIKAEPCAVEAVHLPHFVALQGGH	VDTSLFFQMAHFKEEGMLLEKPARV	1200
1141	POHSWMVAGPLD	IIKAEPCAVEAVHLPHFVALQGGH	VDTSLFFQMAHFKEEGMLLEKPARV	1200
1201	ELHHIVLENP	SPSPGLVLLKMIHNALRFP	IPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR	1260
1201	ELHHIVLENP	SPSPGLVLLKMIHNALRFP	IPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR	1260
1261	KELELCYRSPG	EDQFLFSEFYVCHLGS	GIRLQVKDKDETLLWEALVKPGDLMPATTLIIP	1320
1261	KELELCYRSPG	EDQFLFSEFYVCHLGS	GIRLQVKDKDETLLWEALVKPGDLMPATTLIIP	1320
1321	ARIAPVSPD	ADAPQLLHFVDQYREQLIAR	VTSEVVLKHLGQVLSQEQYERVLAENTRPS	1380
1321	ARIAPVSPD	ADAPQLLHFVDQYREQLIAR	VTSEVVLKHLGQVLSQEQYERVLAENTRPS	1380
1381	QMRKLFSL	SQSMDRCKDGLYQALKETHPHL	MELWEKSGKGLPLSS	1429
1381	QMRKLFSL	SQSMDRCKDGLYQALKETHPHL	MELWEKSGKGLPLSS	1429
RESULT 6				
ABG71633				
ID	ABG71633 standard; Protein; 1429 AA.			
AC	ABG71633;			
XX	10-JAN-2003 (first entry)			
XX	Human caspase recruitment domain-7 (CARD-7).			
DE	Human; caspase activity; caspase recruitment domain-7; CARD-7;			
KW	caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;			
KW	pseudICE; ICEBERG; cell growth; cell death; inflammation;			
KW	apoptosis; caspase activation; cancer; follicular lymphoma;			
KW	leukemia; melanoma; colon cancer; lung carcinoma; viral infection;			
KW	autoimmune disease; systemic lupus erythematosus; reactive arthritis;			
KW	human immunodeficiency virus infection; HIV infection; ALS;			
KW	Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;			
KW	myocardial infarction; stroke; inflammatory disorder; Crohn's disease;			
KW	insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;			
KW	graft rejection; allergic rhinitis; food allergy; conjunctivitis;			
KW	glomerular nephritis; cytostatic; virucide; immunosuppressive;			
XX	dermatological; nephrotropic; neuroprotective; cardiac.			
XX	Homo sapiens.			
OS	US2002128198-A1.			
PN	12-SEP-2002.			
XX	27-NOV-2001; 2001US-0996617.			
XX	28-JUN-1999; 99US-0340620.			
PR	27-OCT-1999; 99US-0428252.			
PR	15-AUG-2001; 2001US-0931071.			
XX	(BERT/) BERTIN J.			
XX	Bertin J;			
PI	WPI; 2003-028967/02.			
DR	N-PSDB; ABS56030.			
DR				

XX Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
PT CARD-7 and CARD-5 in presence of test compound, measuring their
PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
PT altered
XX
PS Disclosure; Fig 1; 43pp; English.
PS
XX The present invention relates to methods of identifying compounds
CC that regulate caspase activity using caspase recruitment domain-7
CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
CC method for identifying a compound that modulates the interaction
CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
CC converting enzyme (pseudolice) or ICEBERG is disclosed. CARD-7 and
CC CARD-8 molecules are useful as modulating agents in regulating a
CC variety of cellular processes including cell growth, cell death, and
CC inflammation. The methods of the invention are useful for identifying
CC compounds that have the ability to increase/decrease apoptosis, or
CC comprise the ability to induce caspase activation. The methods are
CC useful for treating a disorder associated with inappropriate apoptosis
CC or inappropriate inflammation. The methods are useful for treating
CC disorders associated with an undesirably low rate of apoptosis such
CC as cancer (preferably follicular lymphoma, chronic myelogenous
CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
CC infections, autoimmune diseases caused by low levels of apoptosis
CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
CC and arthritis). The methods are also useful for treating disorders with
CC undesirably high rates of apoptosis such as human immunodeficiency
CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
CC muscular atrophy, various forms of cerebellar degeneration, anaemia
CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
CC myelodysplastic syndromes, myocardial infarction, stroke, and
CC various inflammatory disorders (e.g. Crohn's disease, reactive
CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
CC psoriasis, graft rejection, allergic rhinitis, food allergies,
CC conjunctivitis, glomerular nephritis, etc). The present sequence
CC represents human CARD-7.
XX
XX Sequence 1429 AA;
SQ

Query Match 100.0%; Score 7534; DB 24; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWRLACYLEFLKKEELKEFQALLANKAHSRSSSGTETPAQPEKTSQMEVASYLVAQ 60
DB 1 MAGGAWRLACYLEFLKKEELKEFQALLANKAHSRSSSGTETPAQPEKTSQMEVASYLVAQ 60
QY 61 YGEQAWDLALHTWQWGLRSILCAQAEAGHSPSPYPSPHLSQPSQPTSTAVLMPW 120
DB 61 YGEQAWDLALHTWQWGLRSILCAQAEAGHSPSPYPSPHLSQPSQPTSTAVLMPW 120
QY 121 IHLPAGCTQSGERRVRLQPDTSGRWRREISASLLYQALPSPDHESQSPNAPTST 180
DB 121 IHLPAGCTQSGERRVRLQPDTSGRWRREISASLLYQALPSPDHESQSPNAPTST 180
QY 181 AVLGSWSPQPSLAPREOAGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT 240
DB 181 AVLGSWSPQPSLAPREOAGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT 240
QY 241 PQAHTSLQPHHPWPSPVRESLCSWPKWNEDEKFTQLLLQPHPRSDPLVKRSW 300
DB 241 PQAHTSLQPHHPWPSPVRESLCSWPKWNEDEKFTQLLLQPHPRSDPLVKRSW 300
QY 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGR 360
DB 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGR 360
QY 361 FOHVFFSCRELAAQSKVSLAEILGDKGTATAPIRQILSRPERLLFILDGVDGFWLQ 420
DB 361 FOHVFFSCRELAAQSKVSLAEILGDKGTATAPIRQILSRPERLLFILDGVDGFWLQ 420

RESULT 7
AA72669
ID AA72669 standard; Protein; 1473 AA.
XX

421 EPSSSCLHWQSQPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARVVEVLG 480
DB EPSSSCLHWQSQPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARVVEVLG 480
QY PSESRKRYFYRYFTDERQAIAPRLVKSNEKELWALCLVPVWSWLAECTCLQOQMKRKEKL 540
DB PSESRKRYFYRYFTDERQAIAPRLVKSNEKELWALCLVPVWSWLAECTCLQOQMKRKEKL 540
QY TLTSKTTTTLCHVLAQALQAPLQPLRDLCSLAAEGIWOKKTLTLPSPDDLRKHGLDCAI 600
DB TLTSKTTTTLCHVLAQALQAPLQPLRDLCSLAAEGIWOKKTLTLPSPDDLRKHGLDCAI 600
QY ISTFLKMGILQEHPIPLSYSTIHLCFQEFFAAMSIVLEDEKRGKSHSNCCIIDLEKTLBAY 660
DB ISTFLKMGILQEHPIPLSYSTIHLCFQEFFAAMSIVLEDEKRGKSHSNCCIIDLEKTLBAY 660
QY GIHGLFGASTTRFLGLLSDEGEREMENIFHCRISQGRNLMQWVPSLQLLLOPHSLESILH 720
DB GIHGLFGASTTRFLGLLSDEGEREMENIFHCRISQGRNLMQWVPSLQLLLOPHSLESILH 720
QY CLYETRNKTFITQVMAHFEEMGCMVETDMELLVCTFCIKFSRHVKVKKLQIIEGRQHRSTWS 780
DB CLYETRNKTFITQVMAHFEEMGCMVETDMELLVCTFCIKFSRHVKVKKLQIIEGRQHRSTWS 780
QY PTMVVLFRWVPTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE 840
DB PTMVVLFRWVPTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE 840
QY TRLAGCGITAECDKDLAFGLRANQTLTDLISFNLVTDAGAKHLQBLRQPSCKLQRLQ 900
DB TRLAGCGITAECDKDLAFGLRANQTLTDLISFNLVTDAGAKHLQBLRQPSCKLQRLQ 900
QY LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGRLLCEGLRHPACKLIRGLDQ 960
DB LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGRLLCEGLRHPACKLIRGLDQ 960
QY TTLSDEMQRALAEQEKQPLLIFSRKXPSVMTTEGLDTGEMSNSTSSLRQRRLGSERA 1020
DB TTLSDEMQRALAEQEKQPLLIFSRKXPSVMTTEGLDTGEMSNSTSSLRQRRLGSERA 1020
QY ASHVAQANLKLDDVSKIPPIAIEAESSEPVVVELLVCVPSPASQGDHLTKPLGTDGDFW 1080
DB ASHVAQANLKLDDVSKIPPIAIEAESSEPVVVELLVCVPSPASQGDHLTKPLGTDGDFW 1080
QY GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGGIN 1140
DB GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVMREAVTVEIEFCVWDQFLGGIN 1140
QY PQHSMWVAGPLLDIKAEFGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
DB PQHSMWVAGPLLDIKAEFGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
QY ELHHIVLENPSPGLVLLKMIHNAIRPIVTSVVLVYHRVHPPEVTHLYLIPSDCSIR 1260
DB ELHHIVLENPSPGLVLLKMIHNAIRPIVTSVVLVYHRVHPPEVTHLYLIPSDCSIR 1260
QY KELELCYSPGSDQOLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATLIPP 1320
DB KELELCYSPGSDQOLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATLIPP 1320
QY ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHQVLSQEQYERVLAENTRPS 1380
DB ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLHQVLSQEQYERVLAENTRPS 1380
QY QMRKLFSLQSQWDRCKCKDGLYQALKETHPHILMELWEKSKKGLLPLSS 1429
DB QMRKLFSLQSQWDRCKCKDGLYQALKETHPHILMELWEKSKKGLLPLSS 1429

AC AAY72669;
 XX 31-MAY-2001 (first entry)
 DE Human NB-ARC and CARD containing protein (NAC) beta isoform.
 XX Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 XX Domain 329..547
 FT /label= NB domain
 FT /note= "Nucleotide binding domain, also designated as
 FT NB-ARC domain"
 FT Domain 329..341
 FT /label= Walker A
 FT /note= "Also designated as P-loop"
 FT Domain 406..414
 FT /label= Walker B
 FT Region 809..833
 FT /label= Leucine rich_repeat_region
 FT Region 838..862
 FT /label= Leucine rich_repeat_region
 FT Region 865..890
 FT /label= Leucine rich_repeat_region
 FT Region 895..919
 FT /label= Leucine rich_repeat_region
 FT Region 923..947
 FT /label= Leucine rich_repeat_region
 FT Region 957..987
 FT /note= "This 31 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) and NAC delta isoform
 FT (AAY72671) due to alternative mRNA splicing"
 FT Region 1261..1306
 FT /note= "This 45 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) due to alternative
 FT mRNA splicing"
 FT Domain 1079..1364
 FT /note= "TIM-Barrel-like domain"
 FT Domain 1128..1473
 FT /label= CARD-L
 FT /note= "Caspase-associated recruitment domain"
 FT Domain 1128..1261
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT Domain 1306..1473
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT Domain 1373..1473
 FT /label= CARD
 FT /note= "Caspase-associated recruitment domain"
 XX WO200116170-A2.
 PN 08-MAR-2001.
 XX
 PD
 PD
 PF 01-SEP-2000; 2000WO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX (BURN-) BURNHAM INST.
 PA
 XX Reed JC;
 XX
 XX WPI; 2001-183258/18.
 DR N-PSDB; AAD02760.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment

domains, used to produce polypeptides for screening for modulators of
 apoptosis -
 Claim 15; Page 133-137; 184pp; English.
 XX The present sequence is a human NB-ARC and CARD containing protein
 CC (NAC) beta isoform. NAC beta isoform represents the NAC splice
 CC variant in which both the splice regions are present in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.
 CC Note: This sequence is stated as being the same as that shown as
 CC SEQ ID NO:2 (AAY72711) in figure 1A of the specification. However the
 CC sequences differ at several positions.
 XX SQ Sequence 1473 AA;
 Query Match 99.6%; Score 7502; DB 22; Length 1473;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAKHRSSSGETPAQPEKTSGMVEASLYVAQ 60
 DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAKHRSSSGETPAQPEKTSGMVEASLYVAQ 60
 QY 61 YGEQAWDLALHTWEQMGRLSILCAQAQEGAGHSFYPSPSEPHLGSPSQPTSTAVLMPW 120
 DB 61 YGEQAWDLALHTWEQMGRLSILCAQAQEGAGHSFYPSPSEPHLGSPSQPTSTAVLMPW 120
 QY 121 IHEIPAGTCQSSRRVRLQPLDPTSGRRWRREISASLLYQALPSSPDHESPSQESNPATST 180
 DB 121 IHEIPAGTCQSSRRVRLQPLDPTSGRRWRREISASLLYQALPSSPDHESPSQESNPATST 180
 QY 181 AVLGWSGPPQSLAPREQEAQGTQWPLDETSGIYYTIRERERESKSGKPPMAAVVGT 240
 DB 181 AVLGWSGPPQSLAPREQEAQGTQWPLDETSGIYYTIRERERESKSGKPPMAAVVGT 240
 QY 241 PQAHTSLQPHHPWEPSPVRESLCSSTWPKNEDFNQKFTQLLLQRPSPQDPLVKESW 300
 DB 241 PQAHTSLQPHHPWEPSPVRESLCSSTWPKNEDFNQKFTQLLLQRPSPQDPLVKESW 300
 QY 301 PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
 DB 301 PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
 QY 361 FQHVYFSCRELAQSKVSLAELIGKDGATATPAPRIQLSRPERLLFLDGVDEPGWVLO 420
 DB 361 FQHVYFSCRELAQSKVSLAELIGKDGATATPAPRIQLSRPERLLFLDGVDEPGWVLO 420
 QY 421 EPSSELCLHWSQPADALLGSLCKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
 DB 421 EPSSELCLHWSQPADALLGSLCKTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
 QY 481 FSESSRKEYFYRYFTDERQAIAPFLVKSNKELMALCLVPWVSWLACTCLMQMKRKEKL 540
 DB 481 FSESSRKEYFYRYFTDERQAIAPFLVKSNKELMALCLVPWVSWLACTCLMQMKRKEKL 540
 QY 541 TLTSKTTTTCLHYLAQALQAPLQPLQDLCSLAAGIWKQKTLFSPDDLKKGHLDGAI 600
 DB 541 TLTSKTTTTCLHYLAQALQAPLQPLQDLCSLAAGIWKQKTLFSPDDLKKGHLDGAI 600
 QY 601 ISTFLKMGILQHPPIPLSYSFTHLCQFQFFAAMS VVLEDEKGRGKHSNCIDLEKTLAY 660
 DB 601 ISTFLKMGILQHPPIPLSYSFTHLCQFQFFAAMS VVLEDEKGRGKHSNCIDLEKTLAY 660
 QY 661 GIHGLFGASTTRFLGLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLQPHSLESILH 720

CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.
 CC Note: This sequence is stated as being the same as that shown as
 CC SEQ ID NO:2 (See AAY72669) in page 133-137 of the specification. However
 CC the sequences differ at several positions.
 XX
 SQ

Query Match 99.6%; Score 7502; DB 22; Length 1473;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAGGAWRLACVLEFLKKELEKELFOLLANKAHSRSSSGETPAQPKTSGMEVASVLAQ 60
 DB 1 MAGGAWRLACVLEFLKKELEKELFOLLANKAHSRSSSGETPAQPKTSGMEVASVLAQ 60
 QY 61 YGEQAWDLALHTWEQGLRSLSLCAQAEAGAGHSPSPYSPHGLGSPQPTSTAVLMPW 120
 DB 61 YGEQAWDLALHTWEQGLRSLSLCAQAEAGAGHSPSPYSPHGLGSPQPTSTAVLMPW 120
 QY 121 IHLPAGCTQGSERRVLRQPLDTSGRWRREISASLLYQALPSPDHESPSQSPNAPTST 180
 DB 121 IHLPAGCTQGSERRVLRQPLDTSGRWRREISASLLYQALPSPDHESPSQSPNAPTST 180
 QY 181 AVLGWSPQPSLAPREAGPQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT 240
 DB 181 AVLGWSPQPSLAPREAGPQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT 240
 QY 241 PQAHTSLOPHHPWPSPVRESLCSWPWNEDFNOKFTQLLLQPHRPSODPLVKRW 300
 DB 241 PQAHTSLOPHHPWPSPVRESLCSWPWNEDFNOKFTQLLLQPHRPSODPLVKRW 300
 QY 301 PDYVENRGLHIEIRLDFGPDLTQPRIVILQGAAGIKSTILARQVKEAWGRGLYGR 360
 DB 301 PDYVENRGLHIEIRLDFGPDLTQPRIVILQGAAGIKSTILARQVKEAWGRGLYGR 360
 QY 361 FQHVFFSCRELAQSKVSLAEIIGKDGATPAPIRQILSRPERLLFILDGVDPEGVILQ 420
 DB 361 FQHVFFSCRELAQSKVSLAEIIGKDGATPAPIRQILSRPERLLFILDGVDPEGVILQ 420
 QY 421 EPSSELCLHWSQPADALGSLGKTIILPEASFLITARTALQNLIPSLAQARWVEVLG 480
 DB 421 EPSSELCLHWSQPADALGSLGKTIILPEASFLITARTALQNLIPSLAQARWVEVLG 480
 QY 481 FSESSRKEYFYRTDERQAIRFLVKNKELWALCLVPWVSWLACTCLMQQMKERKEL 540
 DB 481 FSESSRKEYFYRTDERQAIRFLVKNKELWALCLVPWVSWLACTCLMQQMKERKEL 540
 QY 541 TLTSTKTTTLLHYLAQAQPLGQPLDRDCLSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 DB 541 TLTSTKTTTLLHYLAQAQPLGQPLDRDCLSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 QY 601 ISTFLKMGILQEHPIPLSYSFHLCQEPFAAMSVYLEDKGRKHSNCIIDLKLTLEY 660
 DB 601 ISTFLKMGILQEHPIPLSYSFHLCQEPFAAMSVYLEDKGRKHSNCIIDLKLTLEY 660
 QY 661 GIHLFGASTTRFLGLLSDEGEREMENIFHCLRSQGRNLMQWVPSLQLLQPHSLH 720
 DB 661 GIHLFGASTTRFLGLLSDEGEREMENIFHCLRSQGRNLMQWVPSLQLLQPHSLH 720
 QY 721 CLYETRNKTLTQVMAHEFMGCMVETDMELLVCTFCIKPSRHVKKLQIEGQRHSTWS 780
 DB 721 CLYETRNKTLTQVMAHEFMGCMVETDMELLVCTFCIKPSRHVKKLQIEGQRHSTWS 780
 QY 781 PTWVILFRWVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPCLLE 840
 DB 781 PTWVILFRWVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPCLLE 840

QY 841 TURLAGCGLTABDCXDLAFGLRANOTLTLDLSFNVLTDAGAKHLCOBLRQPSCKLORLQ 900
 DB 841 TURLAGCGLTABDCXDLAFGLRANOTLTLDLSFNVLTDAGAKHLCOBLRQPSCKLORLQ 900
 QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960
 DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960
 QY 961 TTLSDEMQRALRQEKFPQLLIFSRKRPVMTPTTEGLDTGEMSNSTSLKQRIGSERA 1020
 DB 961 TTLSDEMQRALRQEKFPQLLIFSRKRPVMTPTTEGLDTGEMSNSTSLKQRIGSERA 1020
 QY 1021 ASHVAQANLKLDDVSKIPIPIAIEAESSEPPVVPVELLCVPSPASQGLHTKPLGTTDDFW 1080
 DB 1021 ASHVAQANLKLDDVSKIPIPIAIEAESSEPPVVPVELLCVPSPASQGLHTKPLGTTDDFW 1080
 QY 1081 GTPGVATEVDKKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVMDQFLGEIN 1140
 DB 1081 GTPGVATEVDKKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVMDQFLGEIN 1140
 QY 1141 PQHSMWAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 DB 1141 PQHSMWAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 QY 1201 ELHHIVLENPSPPLGVLLKMIHNAIRFIPVTSVLLVHRVHPEVTHLYLIPSDCSIR 1260
 DB 1201 ELHHIVLENPSPPLGVLLKMIHNAIRFIPVTSVLLVHRVHPEVTHLYLIPSDCSIR 1260
 QY 1261 K-----EELCYSPGEGDQLF 1276
 DB 1261 K-----EELCYSPGEGDQLF 1276
 QY 1277 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLPATTLIPPARIAVPSPLDAPQLLH 1336
 DB 1277 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLPATTLIPPARIAVPSPLDAPQLLH 1336
 QY 1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEOYERVAENTRPPSQMRKLFSLSQSWDRKC 1396
 DB 1337 FVDQYREQLIARVTSVEVVLDKLHGQVLSQEOYERVAENTRPPSQMRKLFSLSQSWDRKC 1396
 QY 1397 KDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
 DB 1397 KDGLYQALKETHPHLIMELWEKSKKGLPLSS 1429
 QY 1441 KDGLYQALKETHPHLIMELWEKSKKGLPLSS 1473
 DB 1441 KDGLYQALKETHPHLIMELWEKSKKGLPLSS 1473

RESULT 9
 ABG78455
 ID ABG78455 standard; Protein; 1429 AA.
 XX
 AC ABG78455;
 XX AC
 DT 15-NOV-2002 (first entry)
 XX DT
 DE Human caspase recruitment protein 7 protein.
 XX DE
 KW Human; human leucine-rich repeat small intestine I; HLRRS11; asthma;
 KW proliferative disorder; gastrointestinal disorder; renal disorder;
 KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KW immune system; anaemia; human immune deficiency virus; HIV; cancer;
 KW blood coagulation disorder; autoimmune disorder; allergic reaction;
 KW inflammatory condition; cardiovascular disorder; ischaemia;
 KW neurological disorder; infectious disease; cytokine production;
 KW expressed sequence tag; EST.
 XX KW
 OS Homo sapiens.
 XX OS
 PN WO200261086-A2.
 XX PN
 PD 08-AUG-2002.
 XX PD
 PF 20-DEC-2001; 2001WO-US49739.
 XX PF
 PR 22-DEC-2000; 2000US-257774P.
 XX PR

XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX PA Feder J, Ramanathan C, Mintier G;
 XX PI WPI; 2002-619252/66.
 XX DR
 XX New isolated nucleic acid molecules encoding HLRRS11 polypeptides, or
 PT their fragments and homologues, useful for preventing, treating and
 PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
 PT or renal disorders
 XX
 PS Disclosure; Figure 2; 336pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules (I) encoding
 CC human leucine-rich repeat small intestine I (HLRRS11) polypeptides.
 CC The nucleic acid molecules and polypeptides are useful for preventing,
 CC treating and ameliorating medical conditions, such as proliferative,
 CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
 CC related to aberrant calcium regulation or apoptosis modulation, either
 CC directly or indirectly. They are also useful for treating, preventing
 CC and/or diagnosing diseases, disorders and/or conditions of: immune system
 CC by activating or inhibiting the proliferation, differentiation, or
 CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
 CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
 CC virus (HIV) infection, HTLV-BLV infection; blood coagulation disorders,
 CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
 CC myasthenia gravis; asthma or allergic reactions; inflammatory
 CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
 CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
 CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
 CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
 CC pneumonia, or viral, bacterial, and fungal infections. The HLRRS11
 CC polypeptides are useful for modulating cytokine production, antigen
 CC presentation, or other processes such as boosting immune responses.
 CC ABG7454-ABG74874 represent HLRRS11 amino acid sequences and related
 CC amino acid sequences of the invention.
 XX
 SQ Sequence 1429 AA;
 Query Match 99.4%; Score 7488; DB 23; Length 1429;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MAGGAWRLACYLEFLKKELEFOLLANKAHSRSSSGTETPAQPEKTSGMVASYLVAQ 60
 DB 1 MAGGAWRLACYLEFLKKELEFOLLANKAHSRSSSGTETPAQPEKTSGMVASYLVAQ 60
 QY 61 YGEQAWDLALHTWEQMGRLSICAQAGAGHSPFPYSPSPHILGSPSQPTSTAVLMPW 120
 DB 61 YGEQAWDLALHTWEQMGRLSICAQAGAGHSPFPYSPSPHILGSPSQPTSTAVLMPW 120
 QY 121 IHLPAGCTQSGERRVRLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180
 DB 121 IHLPAGCTQSGERRVRLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180
 QY 181 AVLGSWGSPQSLAPREQEAQTPWLDTSIGYYTEIREREREKSEKGRPPAAVWCT 240
 DB 181 AVLGSWGSPQSLAPREQEAQTPWLDTSIGYYTEIREREREKSEKGRPPAAVWCT 240
 QY 241 PQQAHTSLQPHHPHWPSPVRESLCSWPKWKNDFNOKFTQLLLQPHPRSDPLVKRSW 300
 DB 241 PQQAHTSLQPHHPHWPSPVRESLCSWPKWKNDFNOKFTQLLLQPHPRSDPLVKRSW 300
 QY 301 PDYVENRGHLIEIRDLFGGLDTPRIVILQGAAGIKSTLARQVKEAWGRGOLYGR 360
 DB 301 PDYVENRGHLIEIRDLFGGLDTPRIVILQGAAGIKSTLARQVKEAWGRGOLYGR 360
 QY 361 FOHVFFVFCRELAQSKVSLAELIGKDGATATAPIRQILSRPERLLFILDGDEPGWLQ 420
 DB 361 FOHVFFVFCRELAQSKVSLAELIGKDGATATAPIRQILSRPERLLFILDGDEPGWLQ 420
 QY 421 EPSSELCLHWSQOPADALLGSLGKTLTLPASFLITARTTALQNLIPSEARWVEVLG 480

DB 421 EPSSELCLHWSQOPADALLGSLGKTLTLPASFLITARTTALQNLIPSEARWVEVLG 480
 QY 481 PSESSEKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLMOOKRKEKL 540
 DB 481 PSESSEKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLMOOKRKEKL 540
 QY 541 TLTSKTTTTLCLHLYLAQALQAPLGPQLRDLCSLAAEGIWQKTKTLFSPDDLRKHGLDCAI 600
 DB 541 TLTSKTTTTLCLHLYLAQALQAPLGPQLRDLCSLAAEGIWQKTKTLFSPDDLRKHGLDCAI 600
 QY 601 ISTFLKMGILQEHPIPLSYPIHLCFOEFFAAMSIVLEDEKRGHSHNCIIDLKXTLAY 660
 DB 601 ISTFLKMGILQEHPIPLSYPIHLCFOEFFAAMSIVLEDEKRGHSHNCIIDLKXTLAY 660
 QY 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGNLMQWVPSLQLLQPHSLESIH 720
 DB 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGNLMQWVPSLQLLQPHSLESIH 720
 QY 721 CLYETRNKTFITQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS 780
 DB 721 CLYETRNKTFITQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS 780
 QY 781 PMVVLFRWVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
 DB 781 PMVVLFRWVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
 QY 841 TRLAGCGLTAECDKDLAFGRANQTLTELDISFNLTDAGAKHLCORLPSCCKLRLQ 900
 DB 841 TRLAGCGLTAECDKDLAFGRANQTLTELDISFNLTDAGAKHLCORLPSCCKLRLQ 900
 QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRLGLDQ 960
 DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRLGLDQ 960
 QY 961 TTLSDEMQLRALAEQEKPOLLIIFSRKPSVMTPEGLDTGEMNSNSTSLKRQRLGSE 1020
 DB 961 TTLSDEMQLRALAEQEKPOLLIIFSRKPSVMTPEGLDTGEMNSNSTSLKRQRLGSE 1020
 QY 1021 ASHVAQANLKLDDYSKIFPIAIEAESSEPVVVELLCVPSASQGDHHTPLGTDGDFW 1080
 DB 1021 ASHVAQANLKLDDYSKIFPIAIEAESSEPVVVELLCVPSASQGDHHTPLGTDGDFW 1080
 QY 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWNTGLCFVMREAVTVIEIFCVMDQPLGSEIN 1140
 DB 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWNTGLCFVMREAVTVIEIFCVMDQPLGSEIN 1140
 QY 1141 FOHSMWVAGPLDDIKAEFGAVEAVHLPHFVALQGHVDTSLFQMAHFKKEGMLLEKPARV 1200
 DB 1141 FOHSMWVAGPLDDIKAEFGAVEAVHLPHFVALQGHVDTSLFQMAHFKKEGMLLEKPARV 1200
 QY 1201 ELHHIVLENPSFPLGVLLKMIHNALREIPVTSVLLYHRVHPPEVTHLYLIPSDCSIR 1260
 DB 1201 ELHHIVLENPSFPLGVLLKMIHNALREIPVTSVLLYHRVHPPEVTHLYLIPSDCSIR 1260
 QY 1261 KELELCYRSPGEDOLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATLIPP 1320
 DB 1261 KELELCYRSPGEDOLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATLIPP 1320
 QY 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDKLHGQVLSQEQYERVLAENTRPS 1380
 DB 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDKLHGQVLSQEQYERVLAENTRPS 1380
 QY 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
 DB 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429

RESULT 10
 AAO17855
 ID AAO17855 standard; Protein; 1429 AA.
 XX
 AC AAO17855;

XX	20-AUG-2002 (first entry)	QY	301	PDYVEENRGLHIEIRDLFGPGLDTPRIVILQAGIGKSTLARQVKEAWGRGQLYGDR	360
DT		DB	301	PDYVEENRGLHIEIRDLFGPGLDTPRIVILQAGIGKSTLARQVKEAWGRGQLYGDR	360
DE		QY	361	FOHVFYFSCRELAQSKVSVLAELICKOGTATPAPRQILSRPERLLFLDGDVDEPGWVLO	420
XX	Pyrin domain containing protein NALP1-hs.	DB	361	FOHVFYFSCRELAQSKVSVLAELICKOGTATPAPRQILSRPERLLFLDGDVDEPGWVLO	420
KW		QY	421	EPSSSCLHWSQPADALLGSLGKTLIPASFLITARTTALQNLIPSEARVVEVLG	480
KW	Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;	DB	421	EPSSSCLHWSQPADALLGSLGKTLIPASFLITARTTALQNLIPSEARVVEVLG	480
KW	antiarteriosclerotic; anticypriotic; antibacterial; virucide;	QY	481	FSSESRKEYFYRYFTDERQATRAFRVLVSKNKLWALCLVPVWSWLAETCLMOMKREKEL	540
KW	neuroprotective; antiarthritic; antirheumatic; antiasthmatic;	DB	481	FSSESRKEYFYRYFTDERQATRAFRVLVSKNKLWALCLVPVWSWLAETCLMOMKREKEL	540
KW	nephrotropic; osteopathic; neutrotropic; intracellular signal transduction;	QY	541	TUTSKTTTTCLHYLAQALQAOPLGQPLRDLCSLAAEGIWQKKTFLFSPDDLRKHGLDGA	600
KW	inflammation; Alzheimer's disease; infection; psoriasis; asthma;	DB	541	TUTSKTTTTCLHYLAQALQAOPLGQPLRDLCSLAAEGIWQKKTFLFSPDDLRKHGLDGA	600
KW	arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;	QY	601	ISTFLKMGILQEHPIPLSYFHLICFOFFAAMSIVLEDEKGRGKHSNCIIDEKTLLEY	660
XX	osteoarthritis; glomerulonephritis.	DB	601	ISTFLKMGILQEHPIPLSYFHLICFOFFAAMSIVLEDEKGRGKHSNCIIDEKTLLEY	660
OS	Unidentified.	QY	661	GIHGLFGASTTRFLGLGSLDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSLESILH	720
XX		DB	661	GIHGLFGASTTRFLGLGSLDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSLESILH	720
PN	WO200240668-A2.	QY	721	CLYETRNKTFELQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLIQIEGQRHSTWS	780
XX	23-MAY-2002.	DB	721	CLYETRNKTFELQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLIQIEGQRHSTWS	780
XX	30-OCT-2001; 2001WO-BF12545.	QY	781	PTMVVLFWRVPVTDAYWQILFSLVKTNRNLKELDLSGNSLSHSAVSKLCTLRPRCLJE	840
XX	15-NOV-2000; 2000DE-1056687.	DB	781	PTMVVLFWRVPVTDAYWQILFSLVKTNRNLKELDLSGNSLSHSAVSKLCTLRPRCLJE	840
PR	30-NOV-2000; 2000DE-1059595.	QY	841	TURLAGCGLTABDCXDLAFGLRANQTLTDLISFNVLTDAGAKHLICQRLQPSCKLQRIQ	900
XX	(APOT-) APOTECH RES & DEV LTD.	DB	841	TURLAGCGLTABDCXDLAFGLRANQTLTDLISFNVLTDAGAKHLICQRLQPSCKLQRIQ	900
PA	Tschopp J, Martinon F;	QY	901	LVSCLGTSDDCCDLASVLSASPSLKLQNNLDVGVRLLCGLRHACKLIRGLDQ	960
XX		DB	901	LVSCLGTSDDCCDLASVLSASPSLKLQNNLDVGVRLLCGLRHACKLIRGLDQ	960
PI	WPI: 2002-427093/45.	QY	961	TTLSDMRQELRALQEQKPOLLI FSRKPSVMTPTTEGLDTGEMSNSTSLKQRQLGSERA	1020
XX	N-PSDB; AAL47127.	DB	961	TTLSDMRQELRALQEQKPOLLI FSRKPSVMTPTTEGLDTGEMSNSTSLKQRQLGSERA	1020
DR	New DNA encoding protein with pyrin domain, useful for treating	QY	1021	ASHVAQANLKLDVSKIPPIABIESPSEVPVVELLCVPSASQCDLHTKPLGTDGDDFW	1080
PT	diseases involving impaired signal transduction, particularly	DB	1021	ASHVAQANLKLDVSKIPPIABIESPSEVPVVELLCVPSASQCDLHTKPLGTDGDDFW	1080
PT	inflammation, also proteins and antibodies	QY	1081	GPTGVPATEVDKKNLYRVHFPVAGSVRWNTGLCFVWREAVTVEIEFCVWDQFLGEIN	1140
XX	Claim 5; Fig 1; 116pp; German.	DB	1081	GPTGVPATEVDKKNLYRVHFPVAGSVRWNTGLCFVWREAVTVEIEFCVWDQFLGEIN	1140
CC	The present invention relates the DNA and their encoded proteins, where	QY	1141	POHSMWVAGPLDIIKAEPCGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200
CC	the proteins contain at least one PYD (pyrin) domain. These can be used	DB	1141	POHSMWVAGPLDIIKAEPCGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200
CC	to treat diseases associated with impaired intracellular signal	QY	1201	ELHHIVLENPSPSLGVLLKMIHNALRFIPVTSVVLLYHRHPPEEVTFLYLIIPSDCSIR	1260
CC	transduction, particularly inflammation such as psoriasis,	DB	1201	ELHHIVLENPSPSLGVLLKMIHNALRFIPVTSVVLLYHRHPPEEVTFLYLIIPSDCSIR	1260
CC	arteriosclerosis, bacterial or viral infections (particularly meningitis	QY	1261	KELELCYRSPGBDQLFSEPFYVGHLSGIRLQVKDKDETLLVWEALVKPGDLMPATTLIPP	1320
CC	and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,	DB	1261	KELELCYRSPGBDQLFSEPFYVGHLSGIRLQVKDKDETLLVWEALVKPGDLMPATTLIPP	1320
CC	sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's	QY	1321	ARIAVPSLPDAPQLAHFVDQVEQLIARTVTSVEVVLDKLHGQVLSQEQVERVLAENTRPS	1380
CC	and Parkinson's diseases. The present sequence is a protein of the	DB	1321	ARIAVPSLPDAPQLAHFVDQVEQLIARTVTSVEVVLDKLHGQVLSQEQVERVLAENTRPS	1380
XX	invention.	QY	1381	QMRKLFSLSQSMDRCKDGLYQALXETHPLHLMELWEKSGKGLPLSS	1429
SQ	Sequence 1429 AA;				
Query Match 99.4%; Score 7488; DB 23; Length 1429;					
Best Local Similarity 99.4%; Pred. No. 0;					
Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;					
QY	1	MAGGAWRLACYLEFLKKEBELKEFQLLANKAHSRSSSGETPAQPKTSGMEVASVLYAQ	60		
DB	1	MAGGAWRLACYLEFLKKEBELKEFQLLANKAHSRSSSGETPAQPKTSGMEVASVLYAQ	60		
QY	61	YGEQAWDLALHTWEQWGLRSLCAQAEAGHSPSPYSPSEPHLGSPOPTSTAVLMPW	120		
DB	61	YGEQAWDLALHTWEQWGLRSLCAQAEAGHSPSPYSPSEPHLGSPOPTSTAVLMPW	120		
QY	121	IHELPACTQGSRRRLVRLQPLDTSGRWRREISALLYQALPSSPDHSPSQSPNAPTST	180		
DB	121	IHELPACTQGSRRRLVRLQPLDTSGRWRREISALLYQALPSSPDHSPSQSPNAPTST	180		
QY	181	AVLWNGSPQPQSLAPREQAPCTQWPLDTSYIYTEREREREKREKRPWAAVGT	240		
DB	181	AVLWNGSPQPQSLAPREQAPCTQWPLDTSYIYTEREREREKREKRPWAAVGT	240		
QY	241	PPQAHSLQPHHPWPSPRESICSTWPKWNEFNQKFTQLLLQPHPRSQDPLVKRW	300		
DB	241	PPQAHSLQPHHPWPSPRESICSTWPKWNEFNQKFTQLLLQPHPRSQDPLVKRW	300		

DB 1381 QMRKLFSLQSDMRKCKDGLYQALKETHPHLMELWEKSKKGLPLSS 1429

RESULT 11
AAE06758
ID AAE06758 standard; Protein; 1473 AA.

AC AAE06758;

XX 16-OCT-2001 (first entry)

DE Human G-protein coupled receptor-8 (GCREC-8) protein.

XX Human; G-protein coupled receptor-8; GCREC-8; cytostatic; hepatotropic;
KW virucide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective;
KW neotropic; cerebroprotective; hypotensive; tranquiliser; vulnerary;
KW ophthalmological; cell proliferative disorder; actinic keratosis;
KW anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
KW psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; cardiovascular disorder;
KW epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia;
KW anorexia; gastrointestinal disorder; pancreatitis; autoimmune disease;
KW Addison's disease; Crohn's disease; acquired immune deficiency syndrome;
KW AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity;
KW osteoporosis; transgenic animal; gene therapy.

OS Homo sapiens.

XX Key Location/Qualifiers

FF Domain 1216..1237

FT Binding-site 334..341

FT /label= ATP/GTP-binding_site

FT /note= "P-loop"

PN WO200157085-A2.

XX 09-AUG-2001.

XX 01-FEB-2001; 2001WO-US03455.

XX 02-FEB-2000; 2000US-0180093.

XX 11-FEB-2000; 2000US-0182045.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Au-Young J, Yue H;

PI WPI: 2001-488869/53.

DR N-PSDB; AAD12951.

XX Novel isolated human G-protein coupled receptor useful for diagnosing,

PT preventing and treating cell proliferative, neurological,

PT cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic

PT disorders -

XX Claim 1; Page 114-117; 138pp; English.

XX The present sequence is human G-protein coupled receptor-8 (GCREC-8)
CC protein. The present invention relates to GCREC protein and nucleic
CC acids encoding them. GCREC protein, its agonist or antagonist are useful
CC for treating diseases or conditions associated with decreased expression
CC or overexpression of functional GCREC in a patient, where the disorder
CC is selected from cell proliferative disorders such as actinic keratosis,
CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
CC cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
CC disease, Huntington's disease, Parkinson's disease, cardiovascular
CC disorders such as hypertension, vasculitis, varicose veins, gastro-
CC intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
CC pancreatitis, autoimmune/inflammatory disorders such as acquired
CC immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic

CC infections, trauma and metabolic disorders such as diabetes, obesity,
CC osteoporosis. GCREC proteins and their cDNAs are used to assess the
CC effects of exogenous compounds on the expression of GCREC sequences.
CC GCREC cDNA is useful to create knock in humanised animals (pigs) or
CC transgenic animals (mice or rats) to model human disease, for
CC therapeutic or diagnostic purposes, for somatic or germline gene
CC therapy, to generate hybridisation probes useful in mapping the
CC naturally occurring genomic sequence, and in molecular biological
CC techniques.

XX SQ Sequence 1473 AA;

Query Match 99.3%; Score 7484; DB 22; Length 1473;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 1426; Conservative 1; Mismatches 2; Indels 44; Gaps 1;

QY 1 MAGGAWGRACYLEFLKKEELKBFOLLANKAHSRSSGSETPAQPEKTSMEVASYLVAQ 60

DB 1 MAGGAWGRACYLEFLKKEELKBFOLLANKAHSRSSGSETPAQPEKTSMEVASYLVAQ 60

QY 61 YGEQAWDLALHTWQMGLRSICAAQAGAGHSFPYSPSEPHLGSPSQSTAVLMPW 120

DB 61 YGEQAWDLALHTWQMGLRSICAAQAGAGHSFPYSPSEPHLGSPSQSTAVLMPW 120

QY 121 IHELPAGCTQGSERVLRLQPDTSRRWRREISASLLYQALPSSPDHSPSQSPNAPTST 180

DB 121 IHELPAGCTQGSERVLRLQPDTSRRWRREISASLLYQALPSSPDHSPSQSPNAPTST 180

QY 181 AVLGWSGSPQPSLAPREQEAPGTQWPLDETSGIYYTIRERERKSEKGRPPWAAVGT 240

DB 181 AVLGWSGSPQPSLAPREQEAPGTQWPLDETSGIYYTIRERERKSEKGRPPWAAVGT 240

QY 241 PQAHTSLQPHHPWEPVSRESLCSWTWPKNEDFNQKFTQLLLQRPSPQDPVLKESW 300

DB 241 PQAHTSLQPHHPWEPVSRESLCSWTWPKNEDFNQKFTQLLLQRPSPQDPVLKESW 300

QY 301 PDYVEENRGLHIEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGOLYGR 360

DB 301 PDYVEENRGLHIEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGOLYGR 360

QY 361 FQHFVYFSCRELAQSKVVSLAEILGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ 420

DB 361 FQHFVYFSCRELAQSKVVSLAEILGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ 420

QY 421 EPSSELCLHWSQPADALLGSLGKLTILPEASFLITARTTALQNLIPSLQARVVEVLG 480

DB 421 EPSSELCLHWSQPADALLGSLGKLTILPEASFLITARTTALQNLIPSLQARVVEVLG 480

QY 481 PSESRKEYFYRYFTDERQAIKRAFLVKSNEKLMALCLVPWVSWLACTCLMQMKRKEKL 540

DB 481 PSESRKEYFYRYFTDERQAIKRAFLVKSNEKLMALCLVPWVSWLACTCLMQMKRKEKL 540

QY 541 TLTSKTTTTCLHLYLAQALQAPLGPQRLDCLSLAAGIWKQKTLFSPDDLRKHGLDGA 600

DB 541 TLTSKTTTTCLHLYLAQALQAPLGPQRLDCLSLAAGIWKQKTLFSPDDLRKHGLDGA 600

QY 601 ISTFLKMGILQEHPIPLSYFIHLICFQFFAAMSIVLEDEKRGKHSNCIIDLKTLAY 660

DB 601 ISTFLKMGILQEHPIPLSYFIHLICFQFFAAMSIVLEDEKRGKHSNCIIDLKTLAY 660

QY 661 GHGFLGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESILH 720

DB 661 GHGFLGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESILH 720

QY 721 CLYETRNKTLFTQVMAHFEEMGCMVETDWELLVCTFCIKFSRHVKKLQLIIEGRQHRSTWS 780

DB 721 CLYETRNKTLFTQVMAHFEEMGCMVETDWELLVCTFCIKFSRHVKKLQLIIEGRQHRSTWS 780

QY 781 PTMVVLFWRVPVTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVKSICKTLRRRCILLE 840

DB 781 PTMVVLFWRVPVTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVKSICKTLRRRCILLE 840

QY 841 TURLAGCGLTAEDCKDLAFGLRANQTLTDLISFNVLTDAGAKHLQRLQRPSCKLQRLQ 900

Db	841	TURLAGCGLTABDCXDLAFGLRANOTLTEDLSFNVLTDAKAKHLCORLQPSCKLRLQ	900
Qy	901	LVSCGLTSDCCODLASVLSASPSLKDLDQNNLDVGVRLCEGLRHPACKLIRLGLDQ	960
Db	901	LVSCGLTSDCCODLASVLSASPSLKDLDQNNLDVGVRLCEGLRHPACKLIRLGLDQ	960
Qy	961	TTLSDEMRELALBOEKPOLLIFFRRKPSVMTPTGLDTGEMSNSTSLKRLGSGERA	1020
Db	961	TTLSDEMRELALBOEKPOLLIFFRRKPSVMTPTGLDTGEMSNSTSLKRLGSGERA	1020
Qy	1021	ASHVAQANIKLIDVSKIPPIAIAESSPEVVPVELLCVPSASQGLHTKPLGTDDDFW	1080
Db	1021	ASHVAQANIKLIDVSKIPPIAIAESSPEVVPVELLCVPSASQGLHTKPLGTDDDFW	1080
Qy	1081	GPTGPVATEVDKKNLYRVHPFPVAGSYRWPNTGLCFVNRREAVTVIEIFCVWDQFLGEIN	1140
Db	1081	GPTGPVATEVDKKNLYRVHPFPVAGSYRWPNTGLCFVNRREAVTVIEIFCVWDQFLGEIN	1140
Qy	1141	POHSMWVAGPLLDIIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPARV	1200
Db	1141	POHSMWVAGPLLDIIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPARV	1200
Qy	1201	ELHHIVLENPSPGLGVLLKMHNLRFIPVTSVLLYHRVHPPEVTHLYLIPDCSIR	1260
Db	1201	ELHHIVLENPSPGLGVLLKMHNLRFIPVTSVLLYHRVHPPEVTHLYLIPDCSIR	1260
Qy	1261	-----KELELCYRSPGEDOLF	1276
Db	1261	KAIDLEMKFQVRTHKPPPLTPLYMGCRYTVSGSGMLEILPKLELCYRSPGEDOLF	1320
Qy	1277	SEFYVGHGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH	1336
Db	1321	SESYVGHGSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH	1380
Qy	1337	FVDQIREQLIARVTSVEVVLDKLGQVLQSOQYERVLAEINTRPSQWRKLFSLQSQWDRKC	1396
Db	1381	FVDQIREQLIARVTSVEVVLDKLGQVLQSOQYERVLAEINTRPSQWRKLFSLQSQWDRKC	1440
Qy	1397	KDGLYQALKETHPHLLIMELWEKSGKGLPLSS	1429
Db	1441	KDGLYQALKETHPHLLIMELWEKSGKGLPLSS	1473
RESULT 12			
ID	AA72670		
DT	AA72670	standard; Protein; 1397 AA.	
AC	AA72670;		
XX			
DT	31-MAY-2001	(first entry)	
XX			
DE	Human NB-ARC and CARD containing protein (NAC) gamma isoform.		
XX			
KW	Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;		
KW	caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;		
KW	cysteine aspartyl protease; apoptosis; cytokine production;		
KW	cytokine receptor signalling; therapy; inflammatory disorder; sepsis;		
KW	fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	1..956	
FT		/note= "Corresponds to 1-956 residues of human NAC	
FT		beta isoform (AA72669)"	
FT	Region	957..1229	
FT		/note= "Corresponds to 988-1260 residues of human NAC	
FT		beta isoform (AA72669)"	
FT	Region	1230..1397	
FT		/note= "Corresponds to 1306-1473 residues of human NAC	
FT		beta isoform (AA72669)"	
XX			

PN	WO200116170-A2.		
XX			
PD	08-MAR-2001.		
XX			
PF	01-SEP-2000; 2000WO-US24152.		
XX			
PR	01-SEP-1999; 99US-0388221.		
XX	(BURN-) BURNHAM INST.		
PA	Reed JC;		
PI	WPI; 2001-183258/18.		
XX	DR N-PSDB; AAD02761.		
DR			
PT	Novel nucleic acid encoding NB-ARC and caspase associated recruitment		
PT	domains, used to produce polypeptides for screening for modulators of		
PT	apoptosis -		
XX			
PS	Claim 15; Page 143-148; 184pp; English.		
CC	The present sequence is a human NB-ARC and CARD containing protein		
CC	(NAC) gamma isoform. NAC gamma isoform represents the NAC splice		
CC	variant in which both the splice regions are absent in the translated		
CC	polypeptide. NAC protein comprises a nucleotide binding (NB) domain		
CC	(also referred as NB-ARC domain), a caspase-associated recruitment		
CC	domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine		
CC	aspartyl proteases, are principal effectors of apoptosis. CARD containing		
CC	NAC proteins are used for screening modulators that modulates apoptosis,		
CC	cytokine production, cytokine receptor signalling and other cellular		
CC	processes. NAC can act as an immunogen for the production of polyclonal		
CC	and monoclonal antibodies. It can also be used to diagnose and treat		
CC	inflammatory disorders such as sepsis, fibrosis and arthritis and cancer		
CC	pathologies such as adenocarcinomas and leukaemias.		
XX			
SQ	Sequence 1397 AA;		
Query Match			
Best Local Similarity 97.5%; Score 7342; DB 22; Length 1397;			
Matches 1397; Conservative 97.8%; Pred. No. 0;			
Mismatches 0; Indels 32; Gaps 2;			
Qy	1	MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASYLVAQ	60
Db	1	MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASYLVAQ	60
Qy	61	YGEQRAWDLALHTWEQMLRSLCAQAGAGHSPPSPSPHLSGSPQSTSTAVLMPW	120
Db	61	YGEQRAWDLALHTWEQMLRSLCAQAGAGHSPPSPSPHLSGSPQSTSTAVLMPW	120
Qy	121	IHELPAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHESPSQSPNAPTST	180
Db	121	IHELPAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHESPSQSPNAPTST	180
Qy	181	AVLGSWGSPQPSLAPREQEAPGTOWPLDETSGIYTYTEIREREREKSEKGRPPAAVGT	240
Db	181	AVLGSWGSPQPSLAPREQEAPGTOWPLDETSGIYTYTEIREREREKSEKGRPPAAVGT	240
Qy	241	PPQAHTSLQPHHPWEPSPVRESLCSTWPKNEDFNQKFTQLLLQRPSPRSODPLVKRSW	300
Db	241	PPQAHTSLQPHHPWEPSPVRESLCSTWPKNEDFNQKFTQLLLQRPSPRSODPLVKRSW	300
Qy	301	PDYVEENRGHLIEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR	360
Db	301	PDYVEENRGHLIEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR	360
Qy	361	FOHVYFSCRELAQSKVSLAEILKDGDTATPAPIRQILSRPERLLFLIDGVDPEGWVLQ	420
Db	361	FOHVYFSCRELAQSKVSLAEILKDGDTATPAPIRQILSRPERLLFLIDGVDPEGWVLQ	420
Qy	421	EPSSLCILHWSQPADALLGSLCKTILPEASFLITARTTALQNLIPSLQARWVEVLG	480
Db	421	EPSSLCILHWSQPADALLGSLCKTILPEASFLITARTTALQNLIPSLQARWVEVLG	480

QY 481 FSESSKEYFYRYFTDERQAIKRAFLVKNKELMALCLVPWVSWLACTCLMOQKREKEL 540
 DB 481 FSESSKEYFYRYFTDERQAIKRAFLVKNKELMALCLVPWVSWLACTCLMOQKREKEL 540
 QY 541 TLTSTKTTTLCYLAQAALQAPGLQRLDCLSLAAGIWKTKLPSDDLKRGHLDGAI 600
 DB 541 TLTSTKTTTLCYLAQAALQAPGLQRLDCLSLAAGIWKTKLPSDDLKRGHLDGAI 600
 QY 601 ISTFLKMGILQEHPIPLSYFHLCOEPFAAMSVYLEDKGRGHSNCIIDLKTLRAY 660
 DB 601 ISTFLKMGILQEHPIPLSYFHLCOEPFAAMSVYLEDKGRGHSNCIIDLKTLRAY 660
 QY 661 GHGLFGASTTFLGLLDEGERENEFHCHLSQGRNLMQVPSLQLLLOPHSLESILH 720
 DB 661 GHGLFGASTTFLGLLDEGERENEFHCHLSQGRNLMQVPSLQLLLOPHSLESILH 720
 QY 721 CLYETRANKTFLQVMAHPEMGCMVETDMELLVCTFCIKFSPRHVKQLQIEGRQHRSTWS 780
 DB 721 CLYETRANKTFLQVMAHPEMGCMVETDMELLVCTFCIKFSPRHVKQLQIEGRQHRSTWS 780
 QY 781 PTMVLFRWVPVTDAYWQILFVSLKVTNRNKLDELDSGNSLSHSAVKSCLKTLRRPRCLLE 840
 DB 781 PTMVLFRWVPVTDAYWQILFVSLKVTNRNKLDELDSGNSLSHSAVKSCLKTLRRPRCLLE 840
 QY 841 TLRLACGGLTAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 DB 841 TLRLACGGLTAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 QY 901 LVSCGLTSCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRGLDQ 960
 DB 901 LVSCGLTSCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRL--- 956
 QY 961 TTLSDEMRELQALEQEKQLILFSRKPSPVTPTEGLDTGEMSNSTSLKQRLGSEBA 1020
 DB 957 -----KPSVMTPTTEGLDTGEMSNSTSLKQRLGSEBA 989
 QY 1021 ASHVAQANLKLDDVSKIFFIAEIAESSPEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1080
 DB 990 ASHVAQANLKLDDVSKIFFIAEIAESSPEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1049
 QY 1081 GPTGPVATEVVDKERNLYRHVPFVAGSYRWPNTGLCFVNRREAVTVEIEFCVWDQFLGEIN 1140
 DB 1050 GPTGPVATEVVDKERNLYRHVPFVAGSYRWPNTGLCFVNRREAVTVEIEFCVWDQFLGEIN 1109
 QY 1141 PQHSMVAGPLDIDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMILLEKPARV 1200
 DB 1110 PQHSMVAGPLDIDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMILLEKPARV 1169
 QY 1201 ELHHIVLENPSPGLVKMLHNLARFIPVTSVLLYHRVHPEEVTFLYLIPSDCSIR 1260
 DB 1170 ELHHIVLENPSPGLVKMLHNLARFIPVTSVLLYHRVHPEEVTFLYLIPSDCSIR 1229
 QY 1261 KELELCYRSPGEDQLFSEFYVGHLSGIRLVQDKKDETLVMEALVKPGDLMPATTLIPP 1320
 DB 1230 -ELELCYRSPGEDQLFSEFYVGHLSGIRLVQDKKDETLVMEALVKPGDLMPATTLIPP 1288
 QY 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLGQVLSQEQYERVLAEINTPS 1380
 DB 1289 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLGQVLSQEQYERVLAEINTPS 1348
 QY 1381 QNRKLFSLSQSDRCKCKDGLYQALKEPHLMELWEKSKKGLPLSS 1429
 DB 1349 QNRKLFSLSQSDRCKCKDGLYQALKEPHLMELWEKSKKGLPLSS 1397

RESULT 13

AAV72671

ID AAY72671 standard; Protein; 1442 AA.

XX AAY72671;

AC AAY72671;

XX 31-MAY-2001 (first.entry)

DT DT

Human NB-ARC and CARD containing protein (NAC) delta isoform.
 Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 cysteine aspartyl protease; apoptosis; cytokine production;
 cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
 Homo sapiens.
 Key Location/Qualifiers
 Region 1..956
 /note= "Corresponds to 1-956 residues of human NAC
 beta isoform (AAY72669)"
 Region 957..1442
 /note= "Corresponds to 988-1473 residues of human NAC
 beta isoform (AAY72669)"

WO200116170-A2.

08-MAR-2001.

01-SEP-2000; 2000WO-US24152.

01-SEP-1999; 99US-0388221.

(BURN-) BURNHAM INST.

Reed JC;

WPI; 2001-183258/18.

N-PSDB; AAD02762.

Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of
 apoptosis -

Claim 15; Page 154-158; 184pp; English.

The present sequence is a human NB-ARC and CARD containing protein
 (NAC) delta isoform. NAC delta isoform represents the NAC splice
 variant in which one of the splice region is absent in the translated
 polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 (also referred as NB-ARC domain), a caspase-associated recruitment
 domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
 aspartyl proteases, are principal effectors of apoptosis. CARD containing
 NAC proteins are used for screening modulators that modulates apoptosis,
 cytokine production, cytokine receptor signalling and other cellular
 processes. NAC can act as an immunogen for the production of polyclonal
 and monoclonal antibodies. It can also be used to diagnose and treat
 inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 pathologies such as adenocarcinomas and leukaemias.

Sequence 1442 AA;

Query Match 97.2%; Score 7325.5; DB 22; Length 1442;

Best Local Similarity 94.9%; Pred. No. 0;

Matches 1398; Conservative 0; Mismatches 0; Indels 75; Gaps 2;

QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKETSGMEVASYLVAQ 60

DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKETSGMEVASYLVAQ 60

QY 61 YGEQAWDLALHTWEQMLRSICAQAEAGHSFPYSPSEPHLGSPQSTSTAVLMPW 120

DB 61 YGEQAWDLALHTWEQMLRSICAQAEAGHSFPYSPSEPHLGSPQSTSTAVLMPW 120

QY 121 THELPAGCTQGSRRVRLQPDTSRRWRREISASLLYQALPSSPDHSPSPNAPTST 180

DB 121 THELPAGCTQGSRRVRLQPDTSRRWRREISASLLYQALPSSPDHSPSPNAPTST 180

QY 181 AVLGWSGSPPOPSLAPREQEAPGTOWPLDETSGIYYTIREREREKSKGPPWAAVVG 240

DB 181 AVLGWSGSPPOPSLAPREQEAPGTOWPLDETSGIYYTIREREREKSKGPPWAAVVG 240

```
Db 181 AVLGWSPQPSLAPREQAEFGTQWPLDETSGIYYTEIRERERSEKGRPPMAVVTG 240
Qy 241 PQAHTSLQPHHPWEPSPRESLCSWTWPKNEDFNQKFTQLLLQORPHPRSDPLVKRSW 300
Db 241 PQAHTSLQPHHPWEPSPRESLCSWTWPKNEDFNQKFTQLLLQORPHPRSDPLVKRSW 300
Qy 301 PYPVENRHLIEIRDLFGGLDTEPRIVILQGAAGIGKSTLARQVKEAMGRGQLYGDR 360
Db 301 PYPVENRHLIEIRDLFGGLDTEPRIVILQGAAGIGKSTLARQVKEAMGRGQLYGDR 360
Qy 361 FOHVYFSCRELAQSKVSLAEIIGKGTATAPARIQILSRPERLILFDLGDVEGWLQ 420
Db 361 FOHVYFSCRELAQSKVSLAEIIGKGTATAPARIQILSRPERLILFDLGDVEGWLQ 420
Qy 421 EFSSELCHWSQPADALLGSLGKTIILPEASFLITARTTALQNLIPSELAQWVEVLG 480
Db 421 EFSSELCHWSQPADALLGSLGKTIILPEASFLITARTTALQNLIPSELAQWVEVLG 480
Qy 481 FSESRKEYFYRYFTDERQAIIRAFRLVKSNNKELWALCLVPVWSWLACTCLMQMKRKEKL 540
Db 481 FSESRKEYFYRYFTDERQAIIRAFRLVKSNNKELWALCLVPVWSWLACTCLMQMKRKEKL 540
Qy 541 TLTSTKTTTTLCHLYLAQALQAPGLQPLRDICSLAAGIWKKTILFSPDILKXGLDGA 600
Db 541 TLTSTKTTTTLCHLYLAQALQAPGLQPLRDICSLAAGIWKKTILFSPDILKXGLDGA 600
Qy 601 ISTFLKMGILQHPPIPLSYFTHLCQEPFAAMSVYLEDEKGRGHSNCIIIDLEKTLBAY 660
Db 601 ISTFLKMGILQHPPIPLSYFTHLCQEPFAAMSVYLEDEKGRGHSNCIIIDLEKTLBAY 660
Qy 661 GIHGLFGASTTFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESIH 720
Db 661 GIHGLFGASTTFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESIH 720
Qy 721 CLYETRNKTLFQWMAHFEMGCMVETDMELLVCTFCIFSRHVKKQLQIEGRQHRSTWS 780
Db 721 CLYETRNKTLFQWMAHFEMGCMVETDMELLVCTFCIFSRHVKKQLQIEGRQHRSTWS 780
Qy 781 PTMVLFVRWVPVTDAYQWILFVSLKTRNLKELDLSGNLSHSAVSKLCKTLRRPRCLLE 840
Db 781 PTMVLFVRWVPVTDAYQWILFVSLKTRNLKELDLSGNLSHSAVSKLCKTLRRPRCLLE 840
Qy 841 TLRLACGLTAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
Db 841 TLRLACGLTAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
Qy 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDGVRLICEGLRHHPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDGVRLICEGLRHHPACKLIRL 956
Qy 961 TTLSDEMRQELRALQEKPKQLLIFSRKPSVMTPTGLDGTGEMSNSTSLKQRLGSE 1020
Db 957 -----KPSVMTPTGLDGTGEMSNSTSLKQRLGSE 989
Qy 1021 ASHVAQANLKLDVSKIFPIAETAEBSSEVPVVELLCVPSASQGLDHTKPLGTDDEFW 1080
Db 990 ASHVAQANLKLDVSKIFPIAETAEBSSEVPVVELLCVPSASQGLDHTKPLGTDDEFW 1049
Qy 1081 GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNGLCFVMREAVTVIEFCVWDQFLEIN 1140
Db 1050 GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNGLCFVMREAVTVIEFCVWDQFLEIN 1109
Qy 1141 PQHSMWVAGPLDIDKAEPCGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMLLEKPARV 1200
Db 1110 PQHSMWVAGPLDIDKAEPCGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEGMLLEKPARV 1169
Qy 1201 ELHHIVLENPSFPLGVLKMIHNAIRLFPVTSWLLYHRVPEEYTFHLYIPSDCSR 1260
Db 1170 ELHHIVLENPSFPLGVLKMIHNAIRLFPVTSWLLYHRVPEEYTFHLYIPSDCSR 1229
Qy 1261 -----KELELCYRSPGSDQLF 1276
Db 1230 KAIDDLKMFQFVRHKPPPLTPLYMGRVTVTSGSGMGLIILPKLELCYRSPGSDQLF 1289
```

```
Qy 1277 SEFYVGHLSGIRLQVXDKKDETLVWEALVKPGDLPATTLLIPARIAVPSPLDAPQLLH 1336
Db 1290 SEFYVGHLSGIRLQVXDKKDETLVWEALVKPGDLPATTLLIPARIAVPSPLDAPQLLH 1349
Qy 1337 FVDQYREQLIARVTSVEVLDKLGQVLSQRYERVAENTRPSQMRKLFSLQSMDRKC 1396
Db 1350 FVDQYREQLIARVTSVEVLDKLGQVLSQRYERVAENTRPSQMRKLFSLQSMDRKC 1409
Qy 1397 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1429
Db 1410 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1442

RESULT 14
AAV72673
ID AAV72673 standard; Protein; 1454 AA.
XX
AC AAV72673;
DT 31-MAY-2001 (first entry)
DE Human NAC beta isoform-CARD-X1 chimeric protein.
XX
KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW cysteine aspartyl protease; apoptosis; cytokine production;
KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW chimeric protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..1078
FT /note= "Corresponds to 1-1078 residues of human NAC
FT beta isoform (AAV72669)"
FT Region 1079..1454
FT /note= "Corresponds to 56-431 residues of human CARD-X
FT protein (AAV72672)"
XX
PN WO200116170-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-US24152.
XX
PR 01-SEP-1999; 99US-0388221.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC;
XX
DR WPI; 2001-183258/18.
XX
N-PSDB; AAD02764.
XX
PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT domains, used to produce polypeptides for screening for modulators of
PT apoptosis -
XX
PS Disclosure; Page 168-172; 184pp; English.
XX
CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC beta isoform-CARD-X1 chimeric protein. NAC protein comprises a
CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC domain. CARD-X protein comprises a caspase-associated recruitment domain
CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC for screening modulators that modulates apoptosis, cytokine production,
CC cytokine receptor signalling and other cellular processes. They can act
CC as an immunogen for the production of polyclonal and monoclonal
CC antibodies. They can also be used to diagnose and treat inflammatory
```


XX
PI Reed JC;
XX
DR WPI; 2001-183258/18.
DR N-PSDB; AAD02765.
XX
PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT domains, used to produce polypeptides for screening for modulators of
PT apoptosis -
XX
PS Disclosure; Page 179-183; 184pp; English.
XX
CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC gamma or delta isoform-CARD-X1 chimeric protein. NAC protein comprises a
CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC domain. CARD-X protein comprises a caspase-associated recruitment domain
CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC for screening modulators that modulates apoptosis, cytokine production,
CC cytokine receptor signalling and other cellular processes. They can act
CC as an immunogen for the production of polyclonal and monoclonal
CC antibodies. They can also be used to diagnose and treat inflammatory
CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
CC such as adenocarcinomas and leukaemias.
XX
SQ Sequence 1424 AA;

Query Match 81.9%; Score 6174; DB 22; Length 1424;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

QY 1 MAGAGWRLACYLEFLKKELEKELFQALLANKAHSRSSGGETPAQPEKTSMEVASYLVAQ 60
DB 1 MAGAGWRLACYLEFLKKELEKELFQALLANKAHSRSSGGETPAQPEKTSMEVASYLVAQ 60

QY 61 YGQRAWDLALHTWEQMLRSQAQEGAGHSPSPYSPSEPHLSPSQPTSTAVLMPW 120
DB 61 YGQRAWDLALHTWEQMLRSQAQEGAGHSPSPYSPSEPHLSPSQPTSTAVLMPW 120

QY 121 IHELPAGCTGSSRRVRLQLPDTSGRWRREISASLLYQALPSSPDHSPQSPNAPTST 180
DB 121 IHELPAGCTGSSRRVRLQLPDTSGRWRREISASLLYQALPSSPDHSPQSPNAPTST 180

QY 181 AVLWGSPPOPLAPREQAPGTQWPLDTSYIYTEREREREKSRPPAAVAVGT 240
DB 181 AVLWGSPPOPLAPREQAPGTQWPLDTSYIYTEREREREKSRPPAAVAVGT 240

QY 241 PPOAHTSLQPHHPWEPSPVRESLCSTWPMKNEFNQKFTQLLLQRPSPRSDPLVKRW 300
DB 241 PPOAHTSLQPHHPWEPSPVRESLCSTWPMKNEFNQKFTQLLLQRPSPRSDPLVKRW 300

QY 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQAGAGIKSTLARQVKEAWRGQLYGR 360
DB 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQAGAGIKSTLARQVKEAWRGQLYGR 360

QY 361 FOHVVFSCRELAAQSKVSWAEIIGKDGATAPIRQILSRPERLLFILDGVDPEGWVQ 420
DB 361 FOHVVFSCRELAAQSKVSWAEIIGKDGATAPIRQILSRPERLLFILDGVDPEGWVQ 420

QY 421 EPSSELCLHWSQPPADALLGSLGKTLPEASFLITARTALQNLIPSLQARWVEVLG 480
DB 421 EPSSELCLHWSQPPADALLGSLGKTLPEASFLITARTALQNLIPSLQARWVEVLG 480

QY 481 FSSSRKEYRYFTDERQAIRAFRLVKSNNELWALCLVPWVSWLACTCLMQMKREKL 540
DB 481 FSSSRKEYRYFTDERQAIRAFRLVKSNNELWALCLVPWVSWLACTCLMQMKREKL 540

QY 541 TLTSKTTTLLHYLAQALQAPLGPOLRLCSIAAEGIWOKKTLFSPDRLKHGLDGI 600
DB 541 TLTSKTTTLLHYLAQALQAPLGPOLRLCSIAAEGIWOKKTLFSPDRLKHGLDGI 600

QY 601 ISTFLXMGILQEHPIPLSYFIHLCFQEFFAAMSVLEDEKGRGKHSNCIIDLKLEAY 660

Search completed: January 29, 2004, 13:45:54
Job time : 74.5142 secs

DB 601 ISTFLXMGILQEHPIPLSYFIHLCFQEFFAAMSVLEDEKGRGKHSNCIIDLKLEAY 660
QY 661 GIHGLFGASTTRFLGLGLLSDEGEREMENIFHCRLSQGNLMQWVPSLQQLLQPHSLESIH 720
DB 661 GIHGLFGASTTRFLGLGLLSDEGEREMENIFHCRLSQGNLMQWVPSLQQLLQPHSLESIH 720
QY 721 CLYETRNKTFITQVNAHFEEMCMCVETDMLLVCTFCIKFSRHVKQLQIIEGRQHRSTWS 780
DB 721 CLYETRNKTFITQVNAHFEEMCMCVETDMLLVCTFCIKFSRHVKQLQIIEGRQHRSTWS 780
QY 781 PMVVLFRWVPVTDAYWOILFSVLKVTNKLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
DB 781 PMVVLFRWVPVTDAYWOILFSVLKVTNKLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
QY 841 TLRLAGCGLTABDCDKDLAFGLRANQTLTFLDLSFNVLTDAGAKHLCQRLRQSPCKLRLQ 900
DB 841 TLRLAGCGLTABDCDKDLAFGLRANQTLTFLDLSFNVLTDAGAKHLCQRLRQSPCKLRLQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRGLDQ 960
DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRGLDQ 960
QY 961 TTLSDEMQRQLRALQERKPOLLI FSRKPSVMTPTTEGLDTGEMSNSTSSLKQRLGSERA 1020
DB 961 TTLSDEMQRQLRALQERKPOLLI FSRKPSVMTPTTEGLDTGEMSNSTSSLKQRLGSERA 1020
QY 991 ASHVAQANLLDVSQIPPIAIEAESSEPVVPVELLCVPSPASQGDHHTKPLGTDGDFW 1080
DB 991 ASHVAQANLLDVSQIPPIAIEAESSEPVVPVELLCVPSPASQGDHHTKPLGTDGDFW 1080
QY 1081 GPTGPVATEVDKKNLYRHFPVAGSYRWPNVTGLCFVMREAVTVIEIFCVWDQFLG-BI 1139
DB 1081 GPTGPVATEVDKKNLYRHFPVAGSYRWPNVTGLCFVMREAVTVIEIFCVWDQFLG-BI 1139
QY 1140 NPQHSWVWAGPLDITKAEP-GAVEAVHLPFHVALQGHVDTSLFQMAHFKEEGMLEKPA 1198
DB 1140 NPQHSWVWAGPLDITKAEP-GAVEAVHLPFHVALQGHVDTSLFQMAHFKEEGMLEKPA 1198
QY 1199 RVELHHIVLENPSFPLGVLKMIHNAIRFIPVTSVLLYHVRHVEVTFHLLYIPSDCS 1258
DB 1199 RVELHHIVLENPSFPLGVLKMIHNAIRFIPVTSVLLYHVRHVEVTFHLLYIPSDCS 1258
QY 1277 SEFYVGHLSGTRLOVKOKKDETLWREALVKPCDLMPTTLIPARIAPVSPDLAPQLH 1336
DB 1277 SEFYVGHLSGTRLOVKOKKDETLWREALVKPCDLMPTTLIPARIAPVSPDLAPQLH 1336
QY 1290 SKFYAGOMKEPIQLEITEKRHGTVMDETEKVDLQVAASAPP-----PFSGAA--- 1339
DB 1290 SKFYAGOMKEPIQLEITEKRHGTVMDETEKVDLQVAASAPP-----PFSGAA--- 1339
QY 1337 FVDQYREQLIARVTSVEVVLDKLH-QVLSQBYQYERVAENRPSQMRKLFSLSQSWDRK 1395
DB 1337 FVDQYREQLIARVTSVEVVLDKLH-QVLSQBYQYERVAENRPSQMRKLFSLSQSWDRK 1395
QY 1396 CKDGLYQALKEPHHIMEL 1415
DB 1396 CKDGLYQALKEPHHIMEL 1415
QY 1400 ALDVLFRSISERDPYLVSYL 1419
DB 1400 ALDVLFRSISERDPYLVSYL 1419

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:03:41 ; Search time 1728.98 Seconds
(without alignments)
11475.528 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcctggagaggt.....aggaataagaattacac 5444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5444	100.0	5444	10	US-09-996-617-1
2	5444	100.0	5444	10	US-09-931-071-1
3	5444	100.0	5444	11	US-09-956-712-3
4	5284.2	97.1	6531	11	US-09-956-712-11
5	4882.4	89.7	5100	11	US-09-956-712-12
6	4287	78.7	4287	10	US-09-996-617-5
7	4287	78.7	4287	10	US-09-931-071-5
8	4148	76.2	4422	10	US-09-388-221-1
9	4100	75.3	4200	10	US-09-388-221-3
10	3958	72.7	4332	10	US-09-388-221-5
11	3400.4	62.5	4556	10	US-09-388-221-9
12	3210.4	59.0	4466	10	US-09-388-221-11
13	2572.6	47.3	2657	11	US-09-895-298-22
14	1710.8	31.4	9649	11	US-09-956-712-10
15	694.4	12.8	720	13	US-10-027-632-150437

16	694.4	12.8	720	14	US-10-027-632-150437
17	564	10.4	564	13	US-10-029-386-2806
18	508.8	9.3	725	13	US-10-027-632-20320
19	508.8	9.3	725	14	US-10-027-632-20320
20	489.4	9.0	491	10	US-09-796-692-5036
21	489.4	9.0	491	12	US-10-057-475B-5036
22	489.4	9.0	491	12	US-10-154-884B-5036
23	489.4	9.0	491	15	US-10-040-862-5036
24	484.4	8.9	578	11	US-09-764-891-1745
25	482.8	8.9	487	15	US-10-060-036-3298
26	414.4	7.6	416	15	US-10-060-036-3299
27	399.8	7.3	3186	15	US-10-066-521-17
28	398.2	7.3	3466	12	US-10-108-260A-718
29	398	7.3	3108	11	US-09-965-621-23
30	398	7.3	3108	12	US-10-407-866-23
31	398	7.3	3218	12	US-10-407-866-67
32	369.8	6.8	3102	13	US-10-132-967-6
33	369.8	6.8	3102	14	US-10-127-516-6
34	369.8	6.8	3102	14	US-10-027-629-6
35	369.8	6.8	3857	13	US-10-132-967-4
36	369.8	6.8	3857	14	US-10-127-516-4
37	369.8	6.8	3857	14	US-10-027-629-4
38	340.6	6.3	371	12	US-10-242-535A-20893
39	318.4	5.8	322	10	US-09-796-692-2680
40	318.4	5.8	322	12	US-10-057-475B-2680
41	318.4	5.8	322	12	US-10-154-884B-2680
42	318.4	5.8	322	15	US-10-040-862-2680
43	313.4	5.8	2691	15	US-10-066-521-21
44	292.8	5.4	2524	11	US-09-965-621-58
45	292.8	5.4	2524	12	US-10-407-866-58

ALIGNMENTS

RESULT 1

US-09-996-617-1
; Sequence 1, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4809)
US-09-996-617-1

Query Match 100.0%; Score 5444; DB 10; Length 5444;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGGCGCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGAGCCGGGGTCCACTCT 60

DB 1 GCCCAGGCGCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGAGCCGGGGTCCACTCT 60

QY 61 GGGTCTGAAGCCCAATTCCTGCTGGGTCTCTCCACCCACCTTCTTCAGCCTT 120
|||||

Qy	2281	AGTCCAGATGACCTCAGGAAGCATGGGTAGATGCGGGCCATCATCTCCACCTTCTTTGAAG	2340
Db	2281	AGTCCAGATGACCTCAGGAAGCATGGGTAGATGCGGGCCATCATCTCCACCTTCTTTGAAG	2340
Qy	2341	ATGGGTATTCCTCAAGAGCACCCCATCCCTCTGAGCTTACAGCTTCAATCACCTCTGTTC	2400
Db	2341	ATGGGTATTCCTCAAGAGCACCCCATCCCTCTGAGCTTACAGCTTCAATCACCTCTGTTC	2400
Qy	2401	CAAGAGTCTTTTGACGCAATGTCTCTTGAGGATGAGAAGGGAGGATGAACAT	2460
Db	2401	CAAGAGTCTTTTGACGCAATGTCTCTTGAGGATGAGAAGGGAGGATGAACAT	2460
Qy	2461	TCTAAATTGCATCATAGATTTTGGAAAAGACGCTAGAAGCATATGGAATACATGCGCTGTTT	2520
Db	2461	TCTAAATTGCATCATAGATTTTGGAAAAGACGCTAGAAGCATATGGAATACATGCGCTGTTT	2520
Qy	2521	GGGCACTCAACACACACGCTTTCCTATTTGGGCCTGTTAAGTCAATGAGGGGAGAGAGATG	2580
Db	2521	GGGCACTCAACACACACGCTTTCCTATTTGGGCCTGTTAAGTCAATGAGGGGAGAGAGATG	2580
Qy	2581	GAGAAATCTTTTCACTCCCGGCTGTCTCAGGGAGGAACTGATGCACTGGGTCCCGTCC	2640
Db	2581	GAGAAATCTTTTCACTCCCGGCTGTCTCAGGGAGGAACTGATGCACTGGGTCCCGTCC	2640
Qy	2641	CTGAGCTGTCTGTGACGACACACTCTCTGGAGTCCCTCACTGCTCTGTATGAGAGACTCGG	2700
Db	2641	CTGAGCTGTCTGTGACGACACACTCTCTGGAGTCCCTCACTGCTCTGTATGAGAGACTCGG	2700
Qy	2701	AACAAACGTTCTGACACAAAGTATGGCCCATTTTCGAAAGAAATGGCATGTGTGTAGAA	2760
Db	2701	AACAAACGTTCTGACACAAAGTATGGCCCATTTTCGAAAGAAATGGCATGTGTGTAGAA	2760
Qy	2761	ACAGACATGAGCTCTTAGTGTGCATTTCTGCAATTAATTTACGCGCCACCTGATGTAAG	2820
Db	2761	ACAGACATGAGCTCTTAGTGTGCATTTCTGCAATTAATTTACGCGCCACCTGATGTAAG	2820
Qy	2821	CTTCAGCTGATGAGGCGACGACACAGATCAACATGAGACGCCACCATGTTAGTCTTG	2880
Db	2821	CTTCAGCTGATGAGGCGACGACACAGATCAACATGAGACGCCACCATGTTAGTCTTG	2880
Qy	2881	TTCAGTGGGTCCAGTCAAGATGCCATTTGGCAGATTCCTCTCCGTCCTCAGGTC	2940
Db	2881	TTCAGTGGGTCCAGTCAAGATGCCATTTGGCAGATTCCTCTCCGTCCTCAGGTC	2940
Qy	2941	ACCAGAACTCTGAAGGAGCTTGAAGTGGAAACTCGCTGAGGCCACTCTGCAGTGAAG	3000
Db	2941	ACCAGAACTCTGAAGGAGCTTGAAGTGGAAACTCGCTGAGGCCACTCTGCAGTGAAG	3000
Qy	3001	AGTCTTTTGAAGACCTCTGAGACGCCCTCGTCCCTCTGAGACCCCTGCGTTGGCTGGC	3060
Db	3001	AGTCTTTTGAAGACCTCTGAGACGCCCTCGTCCCTCTGAGACCCCTGCGTTGGCTGGC	3060
Qy	3061	TGTGGCCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCACAGACC	3120
Db	3061	TGTGGCCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCACAGACC	3120
Qy	3121	CTGACCGAGCTGGACCTTCAATGTGCTACCGATGCTGGAGCCAAACCTTTGC	3180
Db	3121	CTGACCGAGCTGGACCTTCAATGTGCTACCGATGCTGGAGCCAAACCTTTGC	3180
Qy	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGACGCTGGTCACTGTGGCCTC	3240
Db	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGACGCTGGTCACTGTGGCCTC	3240
Qy	3241	ACGTTGACTGTGTCAGGACCTGGGCTCTGTGTTAGTGCAGCCGCCAGCTGAAGGAG	3300
Db	3241	ACGTTGACTGTGTCAGGACCTGGGCTCTGTGTTAGTGCAGCCGCCAGCTGAAGGAG	3300
Qy	3301	CTGACCTGAGCAGAGAAACACTGGATGAGCTTGGCGTGGACCTGCTGTGAGGGGCTC	3360
Db	3301	CTGACCTGAGCAGAGAAACACTGGATGAGCTTGGCGTGGACCTGCTGTGAGGGGCTC	3360
Qy	3361	AGGCATCTCGCTGCAAACTCATACGCTGGGGCTGGACACAGACAACTCTGAGTGTAGG	3420

Db 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGGCCCAAGCC 780
QY 781 CAGGAAGGGGACGCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAAACCCACCTGGGG 840
Db 781 CAGGAAGGGGACGCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAAACCCACCTGGGG 840
QY 841 TCTCCAGGCAACCACTCTCCAGCGAGTGTCTAATGCCCTTGGATCATGAAATGGCGGG 900
Db 841 TCTCCAGGCAACCACTCTCCAGCGAGTGTCTAATGCCCTTGGATCATGAAATGGCGGG 900
QY 901 GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGAGCG 960
Db 901 GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGAGCG 960
QY 961 CGGTGAGAGAAATCTCTCGCTCACTCTCTACCAAGCTCTTCCAGCTCTCCAGACCAT 1020
Db 961 CGGTGAGAGAAATCTCTCGCTCACTCTCTACCAAGCTCTTCCAGCTCTCCAGACCAT 1020
QY 1021 GAGTCTCCAGGCGAGGTACCCAAAGCCGCCACATCCACAGAGTGTCTGGGAGCTGG 1080
Db 1021 GAGTCTCCAGGCGAGGTACCCAAAGCCGCCACATCCACAGAGTGTCTGGGAGCTGG 1080
QY 1081 GGATCCCACTCAGCCAGCTAGCACCCAGAGCAGAGGCTCCTCGGACCCAAAGG 1140
Db 1081 GGATCCCACTCAGCCAGCTAGCACCCAGAGCAGAGGCTCCTCGGACCCAAAGG 1140
QY 1141 CCTCTGGATGAACGTGAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA 1200
Db 1141 CCTCTGGATGAACGTGAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA 1200
QY 1201 TCAGAGAAAGGAGGCCCATGGGAGCGGTGGTAGGAAGCCGCCCAAGCGGCACACC 1260
Db 1201 TCAGAGAAAGGAGGCCCATGGGAGCGGTGGTAGGAAGCCGCCCAAGCGGCACACC 1260
QY 1261 AGCTTACAGCCCAACACACCCATGGGAGCCTTCTGTAGAGAGAGCCTGTGTCCACA 1320
Db 1261 AGCTTACAGCCCAACACACCCATGGGAGCCTTCTGTAGAGAGAGCCTGTGTCCACA 1320
QY 1321 TGGCCCTGAAATAATGAGGATTTAAACAAATAATCACAAGCTGTACTTCTACAAAGA 1380
Db 1321 TGGCCCTGAAATAATGAGGATTTAAACAAATAATCACAAGCTGTACTTCTACAAAGA 1380
QY 1381 CCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTATGTGGAGGAG 1440
Db 1381 CCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTATGTGGAGGAG 1440
QY 1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGATACCCAGAA 1500
Db 1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGATACCCAGAA 1500
QY 1501 CCTCGCATAGTCATCTGACAGGGGCTGTGGAAATTTGGGAAGTCAACACTGGCCAGGCAG 1560
Db 1501 CCTCGCATAGTCATCTGACAGGGGCTGTGGAAATTTGGGAAGTCAACACTGGCCAGGCAG 1560
QY 1561 GTGAAGGAAGCCTGGGGAGAGCCAGCTGTGTATGGGACCGCTTCAGCATGTCTTCTAC 1620
Db 1561 GTGAAGGAAGCCTGGGGAGAGCCAGCTGTGTATGGGACCGCTTCAGCATGTCTTCTAC 1620
QY 1621 TTGAGTGCAGAGAGCTGGGCCAGTCCAGGTGGTGAAGTCTCGCTGAGCTCATCGAATA 1680
Db 1621 TTGAGTGCAGAGAGCTGGGCCAGTCCAGGTGGTGAAGTCTCGCTGAGCTCATCGAATA 1680
QY 1681 GATGGGACAGCCACTCGGCTCCCATTTAGACAGATCTGTCTAGGCCAGAGCGGCTGCTC 1740
Db 1681 GATGGGACAGCCACTCGGCTCCCATTTAGACAGATCTGTCTAGGCCAGAGCGGCTGCTC 1740
QY 1741 TTCAATCTCGATGGTGTAGATGAGCCAGATGGGTCTTGAGGAGCCGAGTCTTGAGCTC 1800
Db 1741 TTCAATCTCGATGGTGTAGATGAGCCAGATGGGTCTTGAGGAGCCGAGTCTTGAGCTC 1800
QY 1801 TGTCTGACTGGAGCCAGCCACAGCCGGGATGCACTGCTGGGAGTTTGTGGGAAA 1860
Db 1801 TGTCTGACTGGAGCCAGCCACAGCCGGGATGCACTGCTGGGAGTTTGTGGGAAA 1860

QY 1861 ACTATATTTCCGAGGCAATCCTTCTGTATCAAGGCTCGGACACAGCTCTGAGAACCTC 1920
Db 1861 ACTATATTTCCGAGGCAATCCTTCTGTATCAAGGCTCGGACACAGCTCTGAGAACCTC 1920
QY 1921 ATTCCTTTCTTTGGAGCAGGACGTTGGGTAGAGTCTTGGGGTCTCTGAGTCCAGCAGG 1980
Db 1921 ATTCCTTTCTTTGGAGCAGGACGTTGGGTAGAGTCTTGGGGTCTCTGAGTCCAGCAGG 1980
QY 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAAGCAAGCAATTTAGAGCCTTTAGGTTG 2040
Db 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAAGCAAGCAATTTAGAGCCTTTAGGTTG 2040
QY 2041 GTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCCCCCTGGGTGTCTGGCTGCC 2100
Db 2041 GTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCCCCCTGGGTGTCTGGCTGCC 2100
QY 2101 TGCATTTCCCTGTATGACAGATGAGAGCGGAGGAAACCTCACATGACTTCCAGACC 2160
Db 2101 TGCATTTCCCTGTATGACAGATGAGAGCGGAGGAAACCTCACATGACTTCCAGACC 2160
QY 2161 ACCACAAACCTCTGTCTACATTTACCTTGCACAGGCTCTCAAGCTCAGCCATTTGGACCC 2220
Db 2161 ACCACAAACCTCTGTCTACATTTACCTTGCACAGGCTCTCAAGCTCAGCCATTTGGACCC 2220
QY 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCACTCTGGCAAAAAAGACCTTTTC 2280
Db 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCACTCTGGCAAAAAAGACCTTTTC 2280
QY 2281 AGTCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
Db 2281 AGTCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
QY 2341 ATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCACTCACCTCTGTTTC 2400
Db 2341 ATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCACTCACCTCTGTTTC 2400
QY 2401 CAAGAGTCTTTTGCAGCAATGTCTATGCTTGGAGGATGAGAAGGGAGAGGTAAACAT 2460
Db 2401 CAAGAGTCTTTTGCAGCAATGTCTATGCTTGGAGGATGAGAAGGGAGAGGTAAACAT 2460
QY 2461 TCTAATTCATATGATTTGGAAAAGACGCTAGAAAGCATATGGAATACATGAGCTGTTT 2520
Db 2461 TCTAATTCATATGATTTGGAAAAGACGCTAGAAAGCATATGGAATACATGAGCTGTTT 2520
QY 2521 GGGGCATCAACACACGTTTCTTATTTGGCCCTGTAAAGTATGAGGGGAGAGAGATG 2580
Db 2521 GGGGCATCAACACACGTTTCTTATTTGGCCCTGTAAAGTATGAGGGGAGAGAGATG 2580
QY 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGAGTGGGTCCCGTCC 2640
Db 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGAGTGGGTCCCGTCC 2640
QY 2641 CTGAGCTGCTGCTGAGCACAACCTCTCTGAGTCCCTCACTGCTTGTACAGACTCGG 2700
Db 2641 CTGAGCTGCTGCTGAGCACAACCTCTCTGAGTCCCTCACTGCTTGTACAGACTCGG 2700
QY 2701 AACAAAAGTCTTCTGACACAGTGTATGGCCCAATTTGAGAAATGGGCATGTGTAGAA 2760
Db 2701 AACAAAAGTCTTCTGACACAGTGTATGGCCCAATTTGAGAAATGGGCATGTGTAGAA 2760
QY 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTCAGCCGCCACGTGAAGAAG 2820
Db 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTCAGCCGCCACGTGAAGAAG 2820
QY 2821 CTTTACAGTGAATGAGGCGAGCAGACAGATCAACATGGAGCCCAACCATGTAGTCTTG 2880
Db 2821 CTTTACAGTGAATGAGGCGAGCAGCAGACAGATCAACATGGAGCCCAACCATGTAGTCTTG 2880
QY 2881 TTCAGTGGGTCCAGTCAAGATGCTTATGGCAGATTTCTTCTCGTCTCTCAAGGTC 2940
Db 2881 TTCAGTGGGTCCAGTCAAGATGCTTATGGCAGATTTCTTCTCGTCTCTCAAGGTC 2940

QY	2941	ACCAGAAACCTGAAGAGCTGGACCTAAGTGGAAACTCGCTGAGCACTCTCGAGTGAAG	3000
DB	2941	ACCAGAAACCTGAAGAGCTGGACCTAAGTGGAAACTCGCTGAGCACTCTCGAGTGAAG	3000
QY	3001	AGTCTTTGTAAGACCTTGAGACCGCCCTCGCTGCTCTCTGAGACCTCTGCGGTGGCTGC	3060
DB	3001	AGTCTTTGTAAGACCTTGAGACCGCCCTCGCTGCTCTCTGAGACCTCTGCGGTGGCTGC	3060
QY	3061	TGTGGCTCACAGCTCAGAGACTCAAGGACCTTGCTTTGGGCTGAGAGCCAAACAGACC	3120
DB	3061	TGTGGCTCACAGCTCAGAGACTCAAGGACCTTGCTTTGGGCTGAGAGCCAAACAGACC	3120
QY	3121	CTGACCGAGCTGAGACCTGAGCTTCAATGTGCTCAGGATCTGAGACCAACACCTTTGC	3180
DB	3121	CTGACCGAGCTGAGACCTGAGCTTCAATGTGCTCAGGATCTGAGACCAACACCTTTGC	3180
QY	3181	CAGAGACTGAGACCGGAGCTCAAGCTTACAGGACTGAGCTGTGAGCTGTGSCCTC	3240
DB	3181	CAGAGACTGAGACCGGAGCTCAAGCTTACAGGACTGAGCTGTGAGCTGTGSCCTC	3240
QY	3241	ACGTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCCGAGCTGAAGGAG	3300
DB	3241	ACGTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCCGAGCTGAAGGAG	3300
QY	3301	CTAGACCTGCAGCAGAAACCTTGGATGAGCTTGGCTGCGACTGCTCTGTGAGGGGCTC	3360
DB	3301	CTAGACCTGCAGCAGAAACCTTGGATGAGCTTGGCTGCGACTGCTCTGTGAGGGGCTC	3360
QY	3361	AGGCATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAGACAACTCTGAGTGAAG	3420
DB	3361	AGGCATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAGACAACTCTGAGTGAAG	3420
QY	3421	ATGAGCAGGAACTGAGGGCCCTGGAGCAGGAGAACTCAGCTGCTCATCTTCAGCAGA	3480
DB	3421	ATGAGCAGGAACTGAGGGCCCTGGAGCAGGAGAACTCAGCTGCTCATCTTCAGCAGA	3480
QY	3481	CGGAAACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3540
DB	3481	CGGAAACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3540
QY	3541	ACATCTCTAAGCGGCGAGACTCGGATCAGAGAGGGGCTTCCCATGTTGCTCAG	3600
DB	3541	ACATCTCTAAGCGGCGAGACTCGGATCAGAGAGGGGCTTCCCATGTTGCTCAG	3600
QY	3601	GCTAATCTCAAACTCTGAGCGTGAGCAGATCTTCCCAATGCTCAGATTCAGAGGAA	3660
DB	3601	GCTAATCTCAAACTCTGAGCGTGAGCAGATCTTCCCAATGCTCAGATTCAGAGGAA	3660
QY	3661	AGCTCCCGCAGAGTAGTACCGGTGGAACTCTTGTGCTGCTCTCTGCTCTCAAGGG	3720
DB	3661	AGCTCCCGCAGAGTAGTACCGGTGGAACTCTTGTGCTGCTCTCTGCTCTCAAGGG	3720
QY	3721	GACCTGCATACGAACTTTGGGGACTGACGATCTTCTGGGGCCCAACGCGGCTGTG	3780
DB	3721	GACCTGCATACGAACTTTGGGGACTGACGATCTTCTGGGGCCCAACGCGGCTGTG	3780
QY	3781	GCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCACCTTCCCTGTAGCTGC	3840
DB	3781	GCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCACCTTCCCTGTAGCTGC	3840
QY	3841	TCCTACCGCTGGCCCAACCGGCTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG	3900
DB	3841	TCCTACCGCTGGCCCAACCGGCTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG	3900
QY	3901	ATTGAATCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGACAGCTGGATG	3960
DB	3901	ATTGAATCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGACAGCTGGATG	3960
QY	3961	GTGGCAGGGCTCTGTGACATCAAGGCTGAGCTTGGAGCTGTGAACTGTGACCTC	4020
DB	3961	GTGGCAGGGCTCTGTGACATCAAGGCTGAGCTTGGAGCTGTGAACTGTGACCTC	4020
QY	4021	CCTCACTTTGTGCTCTCCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGGCCAC	4080

DB	4021	CCTCACTTTGTGCTCTCCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGGCCAC	4080
QY	4081	TTTAAAGAGGAGGATGCTCTCTGAGAAAGCAGCCAGGCTGGAGCTGCATCACATAGTT	4140
DB	4081	TTTAAAGAGGAGGATGCTCTCTGAGAAAGCAGCCAGGCTGGAGCTGCATCACATAGTT	4140
QY	4141	CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTCTGAAAAATGATCAATAATCCCTG	4200
DB	4141	CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTCTGAAAAATGATCAATAATCCCTG	4200
QY	4201	CGCTTCATTCCTGCTCCTCTCTGCTGCTGCTTACACCGGCTCCATCTCTGAGAACTC	4260
DB	4201	CGCTTCATTCCTGCTCCTCTCTGCTGCTGCTTACACCGGCTCCATCTCTGAGAACTC	4260
QY	4261	ACCTTCACCTCTACCTGATCCCAAGTCACTGCTCCATTCGGAAGAACTGAGCTCTGC	4320
DB	4261	ACCTTCACCTCTACCTGATCCCAAGTCACTGCTCCATTCGGAAGAACTGAGCTCTGC	4320
QY	4321	TATCGAAGCCCTGGAGAAAGACCTGCTTCTCGGAGTTCTACGTTGGCCACTTGGATCA	4380
DB	4321	TATCGAAGCCCTGGAGAAAGACCTGCTTCTCGGAGTTCTACGTTGGCCACTTGGATCA	4380
QY	4381	GGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTGCTGCTGGAGGCTTGGTG	4440
DB	4381	GGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTGCTGCTGGAGGCTTGGTG	4440
QY	4441	AAACCCAGAGACTCTCATGCTGCACTACTCTGATCCCTCCAGCCCGCATAGCCGCTACT	4500
DB	4441	AAACCCAGAGACTCTCATGCTGCACTACTCTGATCCCTCCAGCCCGCATAGCCGCTACT	4500
QY	4501	TCACCTCTGGAGTCCCGCCAGTTGCTGCACTTTGTGGAACCAAGTATCGAGAGCAGCTGATA	4560
DB	4501	TCACCTCTGGAGTCCCGCCAGTTGCTGCACTTTGTGGAACCAAGTATCGAGAGCAGCTGATA	4560
QY	4561	GCCCAGTGAACATCGGTGGAGTTGCTTGGACAACTGCAATGCAAGAGGCTGAGCCAG	4620
DB	4561	GCCCAGTGAACATCGGTGGAGTTGCTTGGACAACTGCAATGCAAGAGGCTGAGCCAG	4620
QY	4621	GAGCAGTACGAGAGGCTGCTGCTGAGAACACAGAGGCCAGCCAGATCGGAAGCTGTC	4680
DB	4621	GAGCAGTACGAGAGGCTGCTGCTGAGAACACAGAGGCCAGCCAGATCGGAAGCTGTC	4680
QY	4681	AGCTTGAGCCAGTCTGGGACCGGAACTGCAAGATGCACTTACCAAGCCCTGAAGAG	4740
DB	4681	AGCTTGAGCCAGTCTGGGACCGGAACTGCAAGATGCACTTACCAAGCCCTGAAGAG	4740
QY	4741	ACCCATCTCACTCATTTATGGAATCTGGGAGAAAGGAGCAAAAGGACTCTCTGCCA	4800
DB	4741	ACCCATCTCACTCATTTATGGAATCTGGGAGAAAGGAGCAAAAGGACTCTCTGCCA	4800
QY	4801	CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTGAAGCTTGGCTTGGCTGACCC	4860
DB	4801	CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTGAAGCTTGGCTTGGCTGACCC	4860
QY	4861	TTCTTTGGGTCTCAGTTCTTCTCTGCAAAACAGTGTGCCATCTGTTTGGCTTCCAGCA	4920
DB	4861	TTCTTTGGGTCTCAGTTCTTCTCTGCAAAACAGTGTGCCATCTGTTTGGCTTCCAGCA	4920
QY	4921	CTAAAGTAAATGAACTTTGATGATGCTTGTGCTGGCATTTATGCTCATGTCAGGGATG	4980
DB	4921	CTAAAGTAAATGAACTTTGATGATGCTTGTGCTGGCATTTATGCTCATGTCAGGGATG	4980
QY	4981	CCACAGGGGGCCCGAGTCCAGTGCCTTAAACAGATCTCAGGATGTCCATCTCGAGCT	5040
DB	4981	CCACAGGGGGCCCGAGTCCAGTGCCTTAAACAGATCTCAGGATGTCCATCTCGAGCT	5040
QY	5041	GGCAAGACCCCTGACAGCTCATAGAGCTCATCTGTGGGCCACACAGAGCCAGCTTGA	5100
DB	5041	GGCAAGACCCCTGACAGCTCATAGAGCTCATCTGTGGGCCACACAGAGCCAGCTTGA	5100
QY	5101	GCCCTCCGGATCCCATCCAGGGCGAAGAAATAGGAGGACATGGAAACCATTTGCTCTC	5160

```

Db 5101 GCGCTCCGATCCATCCAGGCGCAAGAGGAAATAGGAGGGAATGGAAACCAATTTGCGCTC 5160
QY 5161 TGGCTGTGTACAGAGGTGAGCCCAAAATTTGGGGTTTACGGTGGAGGCCAGTGGATTC 5220
Db 5161 TGGCTGTGTACAGAGGTGAGCCCAAAATTTGGGGTTTACGGTGGAGGCCAGTGGATTC 5220
QY 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCAAACAGAGTAAAGTGGAAAGGATTT 5280
Db 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCAAACAGAGTAAAGTGGAAAGGATTT 5280
QY 5281 ATTACAGAAATAAAGAGATACAGAGTCTTTTAGAATTTGTAGCGGCTTTCCAGTT 5340
Db 5281 ATTACAGAAATAAAGAGATACAGAGTCTTTTAGAATTTGTAGCGGCTTTCCAGTT 5340
QY 5341 TTTTACAGAAAAACCCCTATAAAATTTAAAAATTTTAAATTTTAAAGAAATTTAAAAAAT 5400
Db 5341 TTTTACAGAAAAACCCCTATAAAATTTAAAAATTTTAAATTTTAAAGAAATTTAAAAAAT 5400
QY 5401 ACAAAGAAAGAAATTAAGAAATAAGGAATAAGAGTTACCTAC 5444
Db 5401 ACAAAGAAAGAAATTAAGAAATAAGGAATAAGAGTTACCTAC 5444

RESULT 3
US-09-956-712-3
; Sequence 3, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
; FILE REFERENCE: RTS-0326
; CURRENT APPLICATION NUMBER: US/09/956,712
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 3
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4812)
US-09-956-712-3

Query Match 100.0%; Score 5444; DB 11; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCCAGGGCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGCAGCGCGGGCTCCACTCT 60
Db 1 GCGCCAGGGCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGCAGCGCGGGCTCCACTCT 60
QY 61 GGGTTCTGAAGCCCAATTCCTGCTCTGGGCTCTCCACCCCACTCTTCTCAGCCTT 120
Db 61 GGGTTCTGAAGCCCAATTCCTGCTCTGGGCTCTCCACCCCACTCTTCTCAGCCTT 120
QY 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGAGCCAGGAGGGAAGAATCTGAGGAACA 180
Db 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGAGCCAGGAGGGAAGAATCTGAGGAACA 180
QY 181 CAGAACAGTGAGGCTTGGCCACACCCCACTCCCGTCCACACATCTCCCTCCACCTCAC 240
Db 181 CAGAACAGTGAGGCTTGGCCACACCCCACTCTCCCGTCCACACATCTCCCTCCACCTCAC 240
QY 241 CCTCCCTGCTGCGCTTGGACCCCACTCCAGGACCTCCCTATCAGCTGACTTCTCCAGT 300
Db 241 CCTCCCTGCTGCGCTTGGACCCCACTCCAGGACCTCCCTATCAGCTGACTTCTCCAGT 300
QY 301 GTCTTCGAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360
Db 301 GTCTTCGAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360
QY 361 CGGCGTCTATCTGTAGTGCGCTGGGATTTTATAAACTGGGTTCCGAATGCTGAATAAGA 420

```

```

Db 361 CGGCGTCTATCTGTAGTGCGCTGGGATTTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
QY 421 GACGGTAAGAGCAAGGCAAGGACAGCACTGTTCTTGCCTGCTGATACCTCACCAC 480
Db 421 GACGGTAAGAGCAAGGCAAGGACAGCACTGTTCTTGCCTGCTGATACCTCACCAC 480
QY 481 CTGGAAACATCCCCAGACACCTTAACTCCGGGACAGAGATGGCTGGGGAGCCTGG 540
Db 481 CTGGAAACATCCCCAGACACCTTAACTCCGGGACAGAGATGGCTGGGGAGCCTGG 540
QY 541 GCGCGCTGGCTGTTACTTGGAGTTCTGAGAGGAGGAGCTGAAGAGGTTCCAGCTT 600
Db 541 GCGCGCTGGCTGTTACTTGGAGTTCTGAGAGGAGGAGCTGAAGAGGTTCCAGCTT 600
QY 601 CTGCTCGCAATAAAGCGCACTCCAGAGGCTCTTCCGGGTGAGACACCCGCTCAGCCAGAG 660
Db 601 CTGCTCGCAATAAAGCGCACTCCAGAGGCTCTTCCGGGTGAGACACCCGCTCAGCCAGAG 660
QY 661 AAGCAGTGGCATGGAGTGGCTCTGTAACCTGGTGGCTCAGTATGGGAGCAGCGGGCC 720
Db 661 AAGCAGTGGCATGGAGTGGCTCTGTAACCTGGTGGCTCAGTATGGGAGCAGCGGGCC 720
QY 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCCCCCAAGCC 780
Db 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCCCCCAAGCC 780
QY 781 CAGGAAGGGGAGGCGCACTCTCCCTCACTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
Db 781 CAGGAAGGGGAGGCGCACTCTCCCTCACTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
QY 841 TCTCCAGCCAAACCCCACTCCACCGCAGTGTCTTAAATGCCCTGGATCCATGAATGGCGG 900
Db 841 TCTCCAGCCAAACCCCACTCCACCGCAGTGTCTTAAATGCCCTGGATCCATGAATGGCGG 900
QY 901 GGGTGCACCCAGGGCTCAGAGAGAGGGTTTGAACACAGCTGCTTGCACATCTGGACGC 960
Db 901 GGGTGCACCCAGGGCTCAGAGAGAGGGTTTGAACACAGCTGCTTGCACATCTGGACGC 960
QY 961 CGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTTCTTCCAAAGCTCCCCAGACCAT 1020
Db 961 CGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTTCTTCCAAAGCTCCCCAGACCAT 1020
QY 1021 GAGTCTCCAAAGCAGGAGTCAACCAACGCCCACTCCACAGAGTGTCTGGGAGCTGG 1080
Db 1021 GAGTCTCCAAAGCAGGAGTCAACCAACGCCCACTCCACAGAGTGTCTGGGAGCTGG 1080
QY 1081 GGATCCCACTCAGCCAGCCTAGCACCCAGAGAGGAGGCTCTGGGACCCCAATGG 1140
Db 1081 GGATCCCACTCAGCCAGCCTAGCACCCAGAGAGGAGGCTCTGGGACCCCAATGG 1140
QY 1141 CCTCTGGATGAACCTCAGGAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
Db 1141 CCTCTGGATGAACCTCAGGAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
QY 1201 TCAGAGAAAGCAGGCCGCCCTCAGCGGTGGTAGGAAACGCCCAACAGCGGCACACC 1260
Db 1201 TCAGAGAAAGCAGGCCGCCCTCAGCGGTGGTAGGAAACGCCCAACAGCGGCACACC 1260
QY 1261 AGCCTACAGCCCAACACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA 1320
Db 1261 AGCCTACAGCCCAACACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA 1320
QY 1321 TGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTCTACAAAGA 1380
Db 1321 TGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTCTACAAAGA 1380
QY 1381 CCTCACCCAGAACCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTATGTGGAGGAG 1440
Db 1381 CCTCACCCAGAACCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTATGTGGAGGAG 1440
QY 1441 AATCGAGGACATTTAATTGAGATCAGAGACTTTATTGGCCCAAGGCTGGATACCCAAGAA 1500

```


QY 3661 AGCTCCCGAGAGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCCTGCCTCTCAAGGG 3720
 Db |||||
 QY 3661 AGCTCCCGAGAGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCCTGCCTCTCAAGGG 3720
 Db |||||
 QY 3721 GACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGGCTGTG 3780
 Db |||||
 QY 3721 GACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGGCTGTG 3780
 Db |||||
 QY 3781 GCTACTGAGTGTGACAAAGAAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 Db |||||
 QY 3781 GCTACTGAGTGTGACAAAGAAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 Db |||||
 QY 3841 TCCTACCGTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAAAGCGGTGACCGTTGAG 3900
 Db |||||
 QY 3841 TCCTACCGTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAAAGCGGTGACCGTTGAG 3900
 Db |||||
 QY 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGCACAGCTGGATG 3960
 Db |||||
 QY 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGCACAGCTGGATG 3960
 Db |||||
 QY 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACTTC 4020
 Db |||||
 QY 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACTTC 4020
 Db |||||
 QY 4021 CCTCACTTTGTGCTCTCCAAAGGGGCGCATGTGGACACATCCCTGTTCCAAATGGCCAC 4080
 Db |||||
 QY 4021 CCTCACTTTGTGCTCTCCAAAGGGGCGCATGTGGACACATCCCTGTTCCAAATGGCCAC 4080
 Db |||||
 QY 4081 TTTAAAGAGGAGGGATGCTCTCGAGAGACGACGAGCTGGAGCTGATCACAATGTT 4140
 Db |||||
 QY 4081 TTTAAAGAGGAGGGATGCTCTCGAGAGACGACGAGCTGGAGCTGATCACAATGTT 4140
 Db |||||
 QY 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTGAAATGATCATATGCTGCTG 4200
 Db |||||
 QY 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTGAAATGATCATATGCTGCTG 4200
 Db |||||
 QY 4201 CGCTTCATTCCTGCTACCTCTGTGTTGCTTTTACACCGCTCATCTGAGGAAGTC 4260
 Db |||||
 QY 4201 CGCTTCATTCCTGCTACCTCTGTGTTGCTTTTACACCGCTCATCTGAGGAAGTC 4260
 Db |||||
 QY 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTTCCATTCCGGAAGAACTGGAGCTCTGC 4320
 Db |||||
 QY 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTTCCATTCCGGAAGAACTGGAGCTCTGC 4320
 Db |||||
 QY 4321 TATCGAAGCCCTCGAGAGACCACTGCTTCTCGAGTTCACGAGTTCACGAGTTCAGGATCA 4380
 Db |||||
 QY 4321 TATCGAAGCCCTCGAGAGACCACTGCTTCTCGAGTTCACGAGTTCACGAGTTCAGGATCA 4380
 Db |||||
 QY 4381 GGGATCAGGCTGCAAGTGAAGAACAAGAAAGATGAGACTCTGCTGGAGGCTTTGGTG 4440
 Db |||||
 QY 4381 GGGATCAGGCTGCAAGTGAAGAACAAGAAAGATGAGACTCTGCTGGAGGCTTTGGTG 4440
 Db |||||
 QY 4441 AAACCGAGAGATCTATGCTGCTGCAACTACTCTGATCTCCAGCCGCGCATAGCCGTACCT 4500
 Db |||||
 QY 4441 AAACCGAGAGATCTATGCTGCTGCAACTACTCTGATCTCCAGCCGCGCATAGCCGTACCT 4500
 Db |||||
 QY 4501 TCACCTCTGATGATCCCGCAGTGTGCTGCTTGTGACCACTGATCGAGAGAGCTGATA 4560
 Db |||||
 QY 4501 TCACCTCTGATGATCCCGCAGTGTGCTGCTTGTGACCACTGATCGAGAGAGCTGATA 4560
 Db |||||
 QY 4561 GCCCGAGTGCATCGGTGGAGTTGTCTTGGACAACTGATGAGACAGAGTGTGAGCCAG 4620
 Db |||||
 QY 4561 GCCCGAGTGCATCGGTGGAGTTGTCTTGGACAACTGATGAGACAGAGTGTGAGCCAG 4620
 Db |||||
 QY 4621 GAGAGTACGAGAGGTGTGCTGCTGAGAAACAGAGGCCCAAGCAGATGCGGAAAGCTGTTTC 4680
 Db |||||
 QY 4621 GAGAGTACGAGAGGTGTGCTGCTGAGAAACAGAGGCCCAAGCAGATGCGGAAAGCTGTTTC 4680
 Db |||||
 QY 4681 AGCTTGAGCCAGTCTCTGGGACCGGAGTGAAGATGGAATCTTACCAAGCCCTGAAGGAG 4740
 Db |||||
 QY 4681 AGCTTGAGCCAGTCTCTGGGACCGGAGTGAAGATGGAATCTTACCAAGCCCTGAAGGAG 4740
 Db |||||
 QY 4741 ACCCATCCTCACCTCATTATGAACTCTGCGAGAGGGCAGCAAAAAGGAGCTCCTGCCA 4800
 Db |||||

Db 4741 ACCCATCCTCACCTCATTATGAACTCTGGGAGAGGGCAGCAAAAAGGAGCTCCTGCCA 4800
 QY |||||
 Db 4801 CTCACGAGCTGAAGTATCAACACACGAGCCCTTACCCCTTGAAGTCTTGGCTGACCC 4860
 QY |||||
 Db 4801 CTCACGAGCTGAAGTATCAACACACGAGCCCTTGAAGTCTTGGCTGACCC 4860
 QY |||||
 Db 4861 TTTCTTGGGTCTCAGTCTTCTCTGCAAAACAAAGTGTCCATCTGCTTTCAGCA 4920
 QY |||||
 Db 4861 TTTCTTGGGTCTCAGTCTTCTCTGCAAAACAAAGTGTCCATCTGCTTTCAGCA 4920
 QY |||||
 QY 4921 CTAAAGTAAAGAACTTTGATGATGCTTGTGCGGCAATATGCTCCATGCGAGGATG 4980
 Db |||||
 QY 4921 CTAAAGTAAAGAACTTTGATGATGCTTGTGCGGCAATATGCTCCATGCGAGGATG 4980
 Db |||||
 QY 4981 CCACAGGGGGCCCCAGTCCAGGTGGCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Db |||||
 QY 4981 CCACAGGGGGCCCCAGTCCAGGTGGCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Db |||||
 QY 5041 GGCAAGACCCCTGACAGCTCTATAGAGCTCATCTGCTGGCCACAGCAGCCCAAGCTAGA 5100
 Db |||||
 QY 5041 GGCAAGACCCCTGACAGCTCTATAGAGCTCATCTGCTGGCCACAGCAGCCCAAGCTAGA 5100
 Db |||||
 QY 5101 GGCCTCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTC 5160
 Db |||||
 QY 5101 GGCCTCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTC 5160
 Db |||||
 QY 5161 TGGCTGTGTACAGGCTGAGCCCAAAATTTGGGTTTACGCTGGAGGCGACCTGGATTC 5220
 Db |||||
 QY 5161 TGGCTGTGTACAGGCTGAGCCCAAAATTTGGGTTTACGCTGGAGGCGACCTGGATTC 5220
 Db |||||
 QY 5221 TTTGGCTTTGTACAGGAAGTCTTACAGAGCAAGCCAAACAGAGTAAAGTGGAGGAAGTTT 5280
 Db |||||
 QY 5221 TTTGGCTTTGTACAGGAAGTCTTACAGAGCAAGCCAAACAGAGTAAAGTGGAGGAAGTTT 5280
 Db |||||
 QY 5281 ATTGAGAAATTAAGAGGATATCAGAGCTCTTTAGAAATTTGTCTAGCAGGCTTTCCAGTT 5340
 Db |||||
 QY 5281 ATTGAGAAATTAAGAGGATATCAGAGCTCTTTAGAAATTTGTCTAGCAGGCTTTCCAGTT 5340
 Db |||||
 QY 5341 TTTTACCAGAAAAACCCCTATAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
 Db |||||
 QY 5341 TTTTACCAGAAAAACCCCTATAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
 Db |||||
 QY 5401 ACAAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAAT 5444
 Db |||||
 QY 5401 ACAAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAAT 5444
 Db |||||

RESULT 4

US-09-956-712-11
 ; Sequence 11, Application US/09956712
 ; Publication No. US20030092648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan M. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
 ; FILE REFERENCE: RTS-0326
 ; CURRENT APPLICATION NUMBER: US/09/956,712
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 11
 ; LENGTH: 6531
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 US-09-956-712-11

Query Match 97.1%; Score 5284.2; DB 11; Length 6531;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 5441; Conservative 0; Mismatches 3; Indels 135; Gaps 2;
 QY 1 GCCCGAGGCTCGAGAGTCTTGAAGAAACCTGGGAGCCAGCAGCCCGGGCTTCCACTCT 60
 |||||

Db	1	GCCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGAGCCAGCAGCCGGGGCTCCACTCT	60
Qy	61	GGGTTCTGAAGCCCAATTCCTGCTCTGGGCTCCTCCACCCACCTCTTCTCAGCCTT	120
Db	61		
Qy	121	GCAGCTCAAGGGTGTATCTCAGAGTCCAGGACCCAGGAGAGGAAGAATCTGAGGAACA	180
Db	121	GCAGCTCAAGGGTGTATCTCAGAGTCCAGGACCCAGGAGAGGAAGAATCTGAGGAACA	180
Qy	181	CAGAACAGTGAGGGTGGCCACACCCCATCTCCCGTCAACATCTCCCTCACCCTCAC	240
Db	181	CAGAACAGTGAGGGTGGCCACACCCCATCTCCCGTCAACATCTCCCTCACCCTCAC	240
Qy	241	CCTCCCTGCTGGCCTGGACCCCATCCAGGACCTCCCTATCAGCTGACTTCTCCAGT	300
Db	241	CCTCCCTGCTGGCCTGGACCCCATCCAGGACCTCCCTATCAGCTGACTTCTCCAGT	300
Qy	301	GTCTTGACAGCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTACCACTCCCCCTCTAT	360
Db	301	GTCTTGACAGCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTACCACTCCCCCTCTAT	360
Qy	361	CGGCGTCTATCTGTAGGTGCTTGGGATTTATAAACTGGGTTCGAAATGCTGAATAAGA	420
Db	361	CGGCGTCTATCTGTAGGTGCTTGGGATTTATAAACTGGGTTCGAAATGCTGAATAAGA	420
Qy	421	GAGGTAAAGCCCAAGGCAAGACAGCACTGTCTCTGCTGCTGATACCTCACAC	480
Db	421	GAGGTAAAGCCCAAGGCAAGACAGCACTGTCTCTGCTGCTGATACCTCACAC	480
Qy	481	CTGGGAACATCCCCACAGACACCTCTTAATCTCGGACAGAGATGGCTGGCGAGCCTGG	540
Db	481	CTGGGAACATCCCCACAGACACCTCTTAATCTCGGACAGAGATGGCTGGCGAGCCTGG	540
Qy	541	GGCCGCTCGCCTGTATCTTGGAGTTCCTGAAGAAGGAGAGCTGAAGAGATTCCAGCTT	600
Db	541	GGCCGCTCGCCTGTATCTTGGAGTTCCTGAAGAAGGAGAGCTGAAGAGATTCCAGCTT	600
Qy	601	CTGCTCGCAATAAAGCGCACTCCAGAGCTCTTGGGTGAGACACCCCGCTCAGCCAGAG	660
Db	601	CTGCTCGCAATAAAGCGCACTCCAGAGCTCTTGGGTGAGACACCCCGCTCAGCCAGAG	660
Qy	661	AAGACAGTGGCATGAGGTGGCTCGTACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
Db	661	AAGACAGTGGCATGAGGTGGCTCGTACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
Qy	721	TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
Db	721	TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
Qy	781	CAGGAAGGGCAGGCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG	840
Db	781	CAGGAAGGGCAGGCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG	840
Qy	841	TCCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCCCTGGATCCATGAATTCGCGGG	900
Db	841	TCCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCCCTGGATCCATGAATTCGCGGG	900
Qy	901	GGGTGACACCCAGGGCTCAGAGAGAGGGTTTGGACAGCTGCTGACACATCTGGAGCG	960
Db	901	GGGTGACACCCAGGGCTCAGAGAGAGGGTTTGGACAGCTGCTGACACATCTGGAGCG	960
Qy	961	CGCTGGAGAGAAATCTCTGCTGCTCCTCTACCAAGCTCTTCCAAAGCTCCCGAGCCAT	1020
Db	961	CGCTGGAGAGAAATCTCTGCTGCTCCTCTACCAAGCTCTTCCAAAGCTCCCGAGCCAT	1020
Qy	1021	GAGTCTCCAAAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTGTGGGAGCTGG	1080
Db	1021	GAGTCTCCAAAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTGTGGGAGCTGG	1080
Qy	1081	GGATCCCAACCTCAGCCCAAGCTTAGCACCCAGAGCAGGAGCTCTCGGACCCCAATGG	1140
Db	1081	GGATCCCAACCTCAGCCCAAGCTTAGCACCCAGAGCAGGAGCTCTCGGACCCCAATGG	1140

Qy	1141	CCTCTGGATGAAAACGTGAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
Db	1141	CCTCTGGATGAAAACGTGAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
Qy	1201	TCAGAGAAAGCAGGCCCCCATGGGACGGTGTAGAAACGCCCCACAGCGCACACC	1260
Db	1201	TCAGAGAAAGCAGGCCCCCATGGGACGGTGTAGAAACGCCCCACAGCGCACACC	1260
Qy	1261	AGCCTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA	1320
Db	1261	AGCCTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA	1320
Qy	1321	TGGCCCTCGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTCTACAAA	1380
Db	1321	TGGCCCTCGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTCTACAAA	1380
Qy	1381	CCTCACCCCAGAACCAAGATCCCTGTGTCAGAGAACTGGCCAGGCTGGATACCCAGAA	1440
Db	1381	CCTCACCCCAGAACCAAGATCCCTGTGTCAGAGAACTGGCCAGGCTGGATACCCAGAA	1440
Qy	1441	AATCGAGACATTTAATTTAGATCAGAGACTTATTTGGCCAGGCTGGATACCCAGAA	1500
Db	1441	AATCGAGACATTTAATTTAGATCAGAGACTTATTTGGCCAGGCTGGATACCCAGAA	1500
Qy	1501	CCTCGCATAGTCATCTCAGGGGCTCTCTGGAATTGGGAAGTCAACACTGGCCAGGCGAG	1560
Db	1501	CCTCGCATAGTCATCTCAGGGGCTCTCTGGAATTGGGAAGTCAACACTGGCCAGGCGAG	1560
Qy	1561	GTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGCTTCCAGCATGTCTTCTAC	1620
Db	1561	GTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGCTTCCAGCATGTCTTCTAC	1620
Qy	1621	TTGAGCTCAGAGAGCTGGCCCGAGTCAAAGTGTGTAGTCTCGCTGAGCTCATTCGAAAA	1680
Db	1621	TTGAGCTCAGAGAGCTGGCCCGAGTCAAAGTGTGTAGTCTCGCTGAGCTCATTCGAAAA	1680
Qy	1681	GATGGGACAGCCACTCCGGCTCCCATTAGACAGATCTCTTAGGCCAGAGCGGCTGCTC	1740
Db	1681	GATGGGACAGCCACTCCGGCTCCCATTAGACAGATCTCTTAGGCCAGAGCGGCTGCTC	1740
Qy	1741	TTGATCTCGATGGTGTAGATGAGCAGGATGGGTCTTTGAGGAGCCGAGTCTGAGCTC	1800
Db	1741	TTGATCTCGATGGTGTAGATGAGCAGGATGGGTCTTTGAGGAGCCGAGTCTGAGCTC	1800
Qy	1801	TGTCTGCACTGAGACCCACAGCCGCGGATGCACTGCTGGGAGTGTCTGGGGAAA	1860
Db	1801	TGTCTGCACTGAGACCCACAGCCGCGGATGCACTGCTGGGAGTGTCTGGGGAAA	1860
Qy	1861	ACTATACCTCCGAGGCACTCTTCTGATCAGGCTCGGACCAACAGCTCTGAGAACCTC	1920
Db	1861	ACTATACCTCCGAGGCACTCTTCTGATCAGGCTCGGACCAACAGCTCTGAGAACCTC	1920
Qy	1921	ATTCCTCTTTGGAGCAGGACGTTGGGTAGAGTCTCTGGGGTCTCTGAGTCCAGCAGG	1980
Db	1921	ATTCCTCTTTGGAGCAGGACGTTGGGTAGAGTCTCTGGGGTCTCTGAGTCCAGCAGG	1980
Qy	1981	AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTG	2040
Db	1981	AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTG	2040
Qy	2041	GTCAANTCAAAACAGAGCTCTGGGCCCTGTGTCTTGTGCCCTGGGTCTCTGGCTGGCC	2100
Db	2041	GTCAANTCAAAACAGAGCTCTGGGCCCTGTGTCTTGTGCCCTGGGTCTCTGGCTGGCC	2100
Qy	2101	TGCACTTCCTGATGACAGCAGATGAAGCGGAAGAAAACCTCACATGACTTCCAGACC	2160
Db	2101	TGCACTTCCTGATGACAGCAGATGAAGCGGAAGAAAACCTCACATGACTTCCAGACC	2160
Qy	2161	ACCAACACCTCTGTCTACATTTACCTTCCAGGCTCTCCAGCTCAGCCATTGGGAGCC	2220
Db	2161	ACCAACACCTCTGTCTACATTTACCTTCCAGGCTCTCCAGCTCAGCCATTGGGAGCC	2220

QY 2221 CAGCTCAGAGACCTCTGCTCTCTGCTGCTGAGGGGCAATCGGCAAAAAAGACCCCTTTTC 2280
Db 2221 CAGCTCAGAGACCTCTGCTCTCTGCTGCTGAGGGCAATCGGCAAAAAAGACCCCTTTTC 2280
QY 2281 AGTCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTCTTGAAG 2340
Db 2281 AGTCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTCTTGAAG 2340
QY 2341 ATGGGTATTTCTCAAGAGCACCCCACTCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC 2400
Db 2341 ATGGGTATTTCTCAAGAGCACCCCACTCCCTCTGAGCTACAGCTTCATTCACCTCTGTTTC 2400
QY 2401 CAAAGAGTCTTTGACAGCAATGCTATGCTTGGAGGATGAGAAGGGGAGAGTAAACAT 2460
Db 2401 CAAAGAGTCTTTGACAGCAATGCTATGCTTGGAGGATGAGAAGGGGAGAGTAAACAT 2460
QY 2461 TCTAATTGCATCATATAGATTGGAAGAGCGCTAGAGCATATGGAATACATGGCCTGTTT 2520
Db 2461 TCTAATTGCATCATATAGATTGGAAGAGCGCTAGAGCATATGGAATACATGGCCTGTTT 2520
QY 2521 GGGGCATCAACACACAGCTTTCTATTGGGCTGTTAAAGTGATGAGGGGAGAGAGATG 2580
Db 2521 GGGGCATCAACACACAGCTTTCTATTGGGCTGTTAAAGTGATGAGGGGAGAGAGATG 2580
QY 2581 GAGAACATCTTTCACTGCGGCTGCTCAGGGAGGAACCTGATGAGTGGTCCGCTCC 2640
Db 2581 GAGAACATCTTTCACTGCGGCTGCTCAGGGAGGAACCTGATGAGTGGTCCGCTCC 2640
QY 2641 CTGCAGCTGCTGTCAGGCACACTCTCTGGAGTCCCTCCACTGCTTGACGAGACTCGG 2700
Db 2641 CTGCAGCTGCTGTCAGGCACACTCTCTGGAGTCCCTCCACTGCTTGACGAGACTCGG 2700
QY 2701 AACAAACGTTCTCTGACACAAAGTGATGGGCCATTTTCGAAGAAATGGCATGTGTAGAA 2760
Db 2701 AACAAACGTTCTCTGACACAAAGTGATGGGCCATTTTCGAAGAAATGGCATGTGTAGAA 2760
QY 2761 ACAGACATGGAGCTTTAGTGTGSCATTTCTGCAATTAATTCAGCGCCACCGTGAAGAAG 2820
Db 2761 ACAGACATGGAGCTTTAGTGTGSCATTTCTGCAATTAATTCAGCGCCACCGTGAAGAAG 2820
QY 2821 CTTACGCTGATTGAGGGCAGGACAGACAGATCAACATGAGGCCACCATGTAAGTCCCTG 2880
Db 2821 CTTACGCTGATTGAGGGCAGGACAGACAGATCAACATGAGGCCACCATGTAAGTCCCTG 2880
QY 2881 TTCAGGTGGTCCAGTCAAGATGCTTATGAGCAGATTTCTCTTCCGCTCTCAAGGTC 2940
Db 2881 TTCAGGTGGTCCAGTCAAGATGCTTATGAGCAGATTTCTCTTCCGCTCTCAAGGTC 2940
QY 2941 ACCAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCAGTGAAG 3000
Db 2941 ACCAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCAGTGAAG 3000
QY 3001 AGTCTTTTGAAGACCTTGAGACGCGCTCGTCTGCTGAGACCCCTGGGTGGTGGC 3060
Db 3001 AGTCTTTTGAAGACCTTGAGACGCGCTCGTCTGCTGAGACCCCTGGGTGGTGGC 3060
QY 3061 TGTGGCTCAGACGTGAGACTCAAGGACCTTGGCTTTGGGCTGAGAGCCAAACACCTTTGC 3120
Db 3061 TGTGGCTCAGACGTGAGACTCAAGGACCTTGGCTTTGGGCTGAGAGCCAAACACCTTTGC 3120
QY 3121 CTGACCGAGCTGACCTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTGC 3180
Db 3121 CTGACCGAGCTGACCTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTGC 3180
QY 3181 CAGAGACTGAGACGCGAGCTGCAAGCTACAGGACTGAGGATGCTGAGTGTGCTGTCGCTC 3240
Db 3181 CAGAGACTGAGACGCGAGCTGCAAGCTACAGGACTGAGGATGCTGAGTGTGCTGTCGCTC 3240
QY 3241 ACCTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCCGCTGAAGGAG 3300
Db 3241 ACCTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCCGCTGAAGGAG 3300
QY 3301 CTAGACCTGCAGCAGAAACAACTGGATGACGTTGGGCTGCGACTGCTCTGTGAGGGGCTC 3360

Db 3301 CTAGACCTGCAGCAGAAACAACTGGATGACGTTGGCTGCGACTGCTCTGTGAGGGGCTC 3360
QY 3361 AGGCATCTCTGCTGCAAACTCATACGCTGGGGCTGGACCAACAACTCTGAGTGTATGAG 3420
Db 3361 AGGCATCTCTGCTGCAAACTCATACGCTGGGGCTGGACCAACAACTCTGAGTGTATGAG 3420
QY 3421 ATGAGGCAAGAACTCAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAAGCAGA 3480
Db 3421 ATGAGGCAAGAACTCAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAAGCAGA 3480
QY 3481 CGGAACCAAGTGTATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGATATAGC 3540
Db 3481 CGGAACCAAGTGTATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGATATAGC 3540
QY 3541 ACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCGCTTCCCATCTTGTCTCAG 3600
Db 3541 ACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCGCTTCCCATCTTGTCTCAG 3600
QY 3601 GCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCAAATTTGCTGAGATTGCGAGCAG 3657
Db 3601 GCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCAAATTTGCTGAGATTGCGAGCAG 3657
QY 3658 GAAAGCTCCAGAGAGTAGTACCGGTGGAACTCTTGTGGTGGCTTCTCTGCTGCTCAA 3717
Db 3661 GAAAGCTCCAGAGAGTAGTACCGGTGGAACTCTTGTGGTGGCTTCTCTGCTGCTCAA 3720
QY 3718 GGGGACCTGCATACCAAGCTTTTGGGACTGACGATGACTTCTGGGGCCCCACGGGCT 3777
Db 3721 GGGGACCTGCATACCAAGCTTTTGGGACTGACGATGACTTCTGGGGCCCCACGGGCT 3780
QY 3778 GTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTGTAGCT 3837
Db 3781 GTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTGTAGCT 3840
QY 3838 GGCTCCTACCGTGGCCCAACACGGGTCTCTGCTTGTGATGAGAGAGCGGTGACCGTT 3897
Db 3841 GGCTCCTACCGTGGCCCAACACGGGTCTCTGCTTGTGATGAGAGAGCGGTGACCGTT 3900
QY 3898 GAGATTGAATCTGTGTGGGACAGTTCTTGGGTGAGATCAACCCAGACACAGCTGG 3957
Db 3901 GAGATTGAATCTGTGTGGGACAGTTCTTGGGTGAGATCAACCCAGACACAGCTGG 3960
QY 3958 ATGTGGCAGGGCTCTGCTGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGCAC 4017
Db 3961 ATGTGGCAGGGCTCTGCTGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGCAC 4020
QY 4018 CTCCTCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGGCC 4077
Db 4021 CTCCTCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGGCC 4080
QY 4078 CACTTTAAAGAGGGGATGCTCTGGAGAGCCAGCCAGGCTGAGCTGCATCACATA 4137
Db 4081 CACTTTAAAGAGGGGATGCTCTGGAGAGCCAGCCAGGCTGAGCTGCATCACATA 4140
QY 4138 GTTCTGGAAAAACCCAGCTTCTCCCTTGGAGTCTCTCCTGAAAATGATCATATGCC 4197
Db 4141 GTTCTGGAAAAACCCAGCTTCTCCCTTGGAGTCTCTCCTGAAAATGATCATATGCC 4200
QY 4198 CTGGCTTCATTTCCGCTACCTCTGTGGTGTGTTTTACCAACGGGTGCATCTGTAGGAA 4257
Db 4201 CTGGCTTCATTTCCGCTACCTCTGTGGTGTGTTTTACCAACGGGTGCATCTGTAGGAA 4260
QY 4258 GTCACTTCCACTCTACCTGATCCCAAGTGTGCTTCCATTCGGAAAGCCCATATGAT 4302
Db 4261 GTCACTTCCACTCTACCTGATCCCAAGTGTGCTTCCATTCGGAAAGCCCATATGAT 4320
QY 4303 ----- 4302
Db 4321 CTAGAAATGAATTCAGTTTGTGGAAATCCCAAGCCACCCCGCTGACCCACTTTAT 4380
QY 4303 -----AAG 4305
|||

Db 4381 ATGGGCTGTCTTACACTGTGTCTGGGTCTGGTTACGGGATGCTGAAATACCTCCCAAG 4440
Qy 4306 GAACTGGAGCTCTGCTATCGAAGCCCTGGAGAGACAGCTGTTCTCGAGTGTCTACGTT 4365
Db 4441 GAACTGGAGCTTTGCTATCGAAGCCCTGGAGAGACAGCTGTTCTCGAGTGTCTACGTT 4500
Qy 4366 GGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAGAAAGATGAGACTCTGGTG 4425
Db 4501 GGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAGAAAGATGAGACTCTGGTG 4560
Qy 4426 TGGAGGCTTGGTGAACACAGAGATCTCATGCTTGCATCTACTCTGATCCCTCCAGCC 4485
Db 4561 TGGAGGCTTGGTGAACACAGAGATCTCATGCTTGCATCTACTCTGATCCCTCCAGCC 4620
Qy 4486 CGCATAGCGTACTTCACTCTGGATGCCCGCAGTGTCTGACATTTCTGGACAGTAT 4545
Db 4621 CGCATAGCGTACTTCACTCTGGATGCCCGCAGTGTCTGACATTTCTGGACAGTAT 4680
Qy 4546 CGAGAGCAGTGTATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGACAAACTGCATGA 4605
Db 4681 CGAGAGCAGTGTATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGACAAACTGCATGA 4740
Qy 4606 CAGGTCTGAGCCAGAGCAGTACAGAGAGGTGCTGGTGAACACAGAGGCCCGCCAG 4665
Db 4741 CAGGTCTGAGCCAGAGCAGTACAGAGAGGTGCTGGTGAACACAGAGGCCCGCCAG 4800
Qy 4666 ATGCGAAGCTGTTCACTTGAAGCAGTCTGGGACGGAGTGCATGCAACTCTAC 4725
Db 4801 ATGCGAAGCTGTTCACTTGAAGCAGTCTGGGACGGAGTGCATGCAACTCTAC 4860
Qy 4726 CAAAGCCCTGAAGAGACCCATCTCACTTATGGAATCTTGGGAGAGGGCAGCAAA 4785
Db 4861 CAAAGCCCTGAAGAGACCCATCTCACTTATGGAATCTTGGGAGAGGGCAGCAAA 4920
Qy 4786 AAGGAGTCTCTGCCATCAGCAGCTGAAGTATCAACACAGCCCTTGAACCTTGAAGTCT 4845
Db 4921 AAGGAGTCTCTGCCATCAGCAGCTGAAGTATCAACACAGCCCTTGAACCTTGAAGTCT 4980
Qy 4846 GGTCTTGGCTGACCTTCTGGGTCTCAGTCTTCTCTGCAACCAAGTGGCATCTG 4905
Db 4981 GGTCTTGGCTGACCTTCTGGGTCTCAGTCTTCTCTGCAACCAAGTGGCATCTG 5040
Qy 4906 GTTGTGCTTCCAGCATTAAGTAATGGAATCTTGTATGATGCCCTTCTGGGCAATATGTG 4965
Db 5041 GTTGTGCTTCCAGCATTAAGTAATGGAATCTTGTATGATGCCCTTCTGGGCAATATGTG 5100
Qy 4966 TCCATGCCAGGATGCCACAGGGGGCCCAAGTCCAGTCCAGTCCAGTCTCAGGAA 5025
Db 5101 TCCATGCCAGGATGCCACAGGGGGCCCAAGTCCAGTCCAGTCCAGTCTCAGGAA 5160
Qy 5026 TGTCCATCTGGAGCTGGCAAGACCCCTGAGACCTCATAGAGCCTCATCTGGTGGCCACA 5085
Db 5161 TGTCCATCTGGAGCTGGCAAGACCCCTGAGACCTCATAGAGCCTCATCTGGTGGCCACA 5220
Qy 5086 GCAGCCAAGCTTAGAGCCCTCCGATCCATCCAGCGGCAAGAGAAATAGAGGGACAT 5145
Db 5221 GCAGCCAAGCTTAGAGCCCTCCGATCCATCCAGCGGCAAGAGAAATAGAGGGACAT 5280
Qy 5146 GGAACCATTTGCTCTGGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGTTTCAAGTGGG 5205
Db 5281 GGAACCATTTGCTCTGGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGTTTCAAGTGGG 5340
Qy 5206 AGGCCAGTGGATTTTGGCTTTGTACAGGAAGATCTACAGAGCAAGCAACAGAGTAA 5265
Db 5341 AGGCCAGTGGATTTTGGCTTTGTACAGGAAGATCTACAGAGCAAGCAACAGAGTAA 5400
Qy 5266 AGTGAAGGAAGTTTATTCAGAAAAATAAGGAGTATCAGAGCTCTTTTGAATTTGCTTA 5325
Db 5401 AGTGAAGGAAGTTTATTCAGAAAAATAAGGAGTATCAGAGCTCTTTTGAATTTGCTTA 5460
Qy 5326 GCAGGCTTTCCAGTTTTCAGAAAAACCCCTATAAATTTAAATTTTAACTTAATTT 5385
Db 5461 GCAGGCTTTCCAGTTTTCAGAAAAACCCCTATAAATTTAAATTTTAACTTAATTT 5520

Qy 5386 AAGAATTAAAAAATACAAAAAGAAAAATGAAAAATAAAGGAATAAGAAATTACCTAC 5444
Db 5521 AAGAATTAAAAAATACAAAAAGAAAAATGAAAAATAAAGGAATAAGAAATTACCTAC 5579

RESULT 5
US-09-956-712-12
; Sequence 12, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
; FILE REFERENCE: RTS-0326
; CURRENT APPLICATION NUMBER: US/09/956,712
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 12
; LENGTH: 5100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160) ... (4449)
US-09-956-712-12

Query Match 89.7%; Score 4882.4; DB 11; Length 5100;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 516 GACACAGATGGCTGGCGAGCCTGGGGCGCTGGCTGTTACTTTGGAGTTCTCTCAAGAA 575
Db 133 GAACAAGATGGCTGGCGAGCCTGGGGCGCTGGCTGTTACTTTGGAGTTCTCTCAAGAA 212
Qy 576 GGAGGAGCTGAAGGAGTTTCCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTCTTC 635
Db 213 GGAGGAGCTGAAGGAGTTTCCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTCTTC 272
Qy 636 GGGTCAGACACCCGCTCAGCCAGACAGAACGAGTGGCATGGAGTGGCTCGTACTCGT 695
Db 273 GGGTCAGACACCCGCTCAGCCAGACAGAACGAGTGGCATGGAGTGGCTCGTACTCGT 332
Qy 696 GGCTCAGTATGGGAGCAGCGGGCTCGGACCTAGCCCTCCATACCTGGGAGCAGATGGG 755
Db 333 GGCTCAGTATGGGAGCAGCGGGCTCGGACCTAGCCCTCCATACCTGGGAGCAGATGGG 392
Qy 756 GCTGAGGTCACTGTGCGGCCCAAGCCAGAGAGGGGCGAGCCACTCTCCCTCAATCCCCCTA 815
Db 393 GCTGAGGTCACTGTGCGGCCCAAGCCAGAGAGGGGCGAGCCACTCTCCCTCAATCCCCCTA 452
Qy 816 CAGCCCAAGTGAACCCACCTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTCTAAT 875
Db 453 CAGCCCAAGTGAACCCACCTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTCTAAT 512
Qy 876 GCGCTGGATCCATGAATTTGCGCGGGGTGACCCAGAGGCTCAGAGAGAGGGTTTGGAG 935
Db 513 GCGCTGGATCCATGAATTTGCGCGGGGTGACCCAGAGGCTCAGAGAGAGGGTTTGGAG 572
Qy 936 ACAGCTGCTGACACATCTGGAACGCGCTGGAGAGAAATCTTGCTCACTCTCTACCA 995
Db 573 ACAGCTGCTGACACATCTGGAACGCGCTGGAGAGAAATCTTGCTCACTCTCTACCA 632
Qy 996 AGCTTCTCAGAGTCCCGAGACCATGATCTCCAGCCAGAGGTCAACCAAGCCGCCAC 1055
Db 633 AGCTTCTCAGAGTCCCGAGACCATGATCTCCAGCCAGAGGTCAACCAAGCCGCCAC 692
Qy 1056 ATCCACAGCAGTCTGGGGAGCTGGGGATCCCGACCTCAGCCAGCCTAGCAGACCCAGAGA 1115
Db 693 ATCCACAGCAGTCTGGGGAGCTGGGGATCCCGACCTCAGCCAGCCTAGCAGACCCAGAGA 752
Qy 1116 GCAGGAGGCTCCTGGGACCCCAATGGCCTCTGGATGAACAGTCAAGAAATTTACTACACAGA 1175

QY 3336 CBTGCGAGTCTCTGTGAGGGGCTCAGGCATCTCTGCTGCAAACTCATACGCTTGGGCT 3395
 Db 2973 CBTGCGAGTCTCTGTGAGGGGCTCAGGCATCTCTGCTGCAAACTCATACGCTTGGGCT 3032
 QY 3396 GGAACAGACAACTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAA 3455
 Db 3033 GGAACAGACAACTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAA 3092
 QY 3456 ACCTGAGTCTCATCTTCAGCAGACGGAAACCAAGTGATGAGCCCTACTGAGGGCCCT 3515
 Db 3093 GCTCAGTCTCATCTTCAGCAGACGGAAACCAAGTGATGAGCCCTACTGAGGGCCCT 3152
 QY 3516 GGATACGGAGAGATGAGTAATAGCACATCTCCTCACTCAAGCGGCAGAGACTCGGATCAGA 3575
 Db 3153 GGATACGGAGAGATGAGTAATAGCACATCTCCTCACTCAAGCGGCAGAGACTCGGATCAGA 3212
 QY 3576 GAGGGCGCTTCCATGTCTCAGGCTAATCTCAAACTCTTGAGCGTGAGCAAGATCTT 3635
 Db 3213 GAGGGCGCTTCCATGTCTCAGGCTAATCTCAAACTCTTGAGCGTGAGCAAGATCTT 3272
 QY 3636 CCCAATTGCTGAGATTCAGAGAGAGCTCCCGAGGCTAGTACCGTGGAAGCTCTTGTG 3695
 Db 3273 CCCAATTGCTGAGATTCAGAGAGAGCTCCCGAGGCTAGTACCGTGGAAGCTCTTGTG 3332
 QY 3696 CGTGCCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGACTGACGATCA 3755
 Db 3333 CGTGCCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGACTGACGATCA 3392
 QY 3756 CTTCTGGGGCCCAACGGGCGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTA 3815
 Db 3393 CTTCTGGGGCCCAACGGGCGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTA 3452
 QY 3816 CCGAGTTCACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACAGGGTCTCTGCTTGT 3875
 Db 3453 CCGAGTTCACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACAGGGTCTCTGCTTGT 3512
 QY 3876 GATGAGAGAGCGGTGACCGTTGAGATTCGAATTCGTGTGGGACCAAGTCTCTGGGTGA 3935
 Db 3513 GGTGAGAGAGCGGTGACCGTTGAGATTCGAATTCGTGTGGGACCAAGTCTCTGGGTGA 3572
 QY 3936 GATCAACCAACAGCAGCTGAGTGTGAGGCGCTCTGCTGGAACATCAAGGCTGAGCC 3995
 Db 3573 GATCAACCAACAGCAGCTGAGTGTGAGGCGCTCTGCTGGAACATCAAGGCTGAGCC 3632
 QY 3996 TGGAGTGTGGAGCTGCACTCCCTCCTCCTTGTGGCTCTCCAGGGGGCCATGTGA 4055
 Db 3633 TGGAGCGGTGGAGCTGTGCACTCCCTCCTCCTTGTGGCTCTCCAGGGGGCCATGTGA 3692
 QY 4056 CACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCCTCGAGAGCCAGC 4115
 Db 3693 CACATCCCTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCCTCGAGAGCCAGC 3752
 QY 4116 CAGGTGGAGCTGCATCACAATAGTCTGGAAACCCCAAGCTTCTCCCTTGGAGTCT 4175
 Db 3753 CAGGTGGAGCTGCATCACAATAGTCTGGAAACCCCAAGCTTCTCCCTTGGAGTCT 3812
 QY 4176 CTTGAAAATGATCCATAAATGCTGCTGCTTCACTTCCCGTCACTCTGTGGTGTGCTTTA 4235
 Db 3813 CTTGAAAATGATCCATAAATGCTGCTGCTTCACTTCCCGTCACTCTGTGGTGTGCTTTA 3872
 QY 4236 CCACCGCTCCATCTCAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTC 4295
 Db 3873 CCACCGCTCCATCTCAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTC 3932
 QY 4296 CATTCGGAAGGAATCGAGCTCTGCTPATCGAAGCCCTGGAGAGACCAAGCTGTCTCGGA 4355
 Db 3933 CATTCGGAAGGAATCGAGCTCTGCTPATCGAAGCCCTGGAGAGACCAAGCTGTCTCGGA 3992
 QY 4356 GTTCTACGTTGGCACTTGGGATCAGGATCAGGCTGCAAGTGAAGCAAGAAAGATCA 4415
 Db 3993 GTTCTACGTTGGCACTTGGGATCAGGATCAGGCTGCAAGTGAAGCAAGAAAGATCA 4052
 QY 4416 GACTCTGGTGTGGAGGCGCTTGGTGAAACCAAGGAGATCTCATGCTGCAACTCTGAT 4475

RESULT 6

4053 GACTCTGGTGTGGAGGCGCTTGGTGAAACCCAGGAGATCTCATGCTGCAACTACTCTGAT 4112
 QY 4476 CCCTCCAGCCCGCATAGCCGTACCTTCACCTCTGATGCCCGCAGTTGCTGCACCTTGT 4535
 Db 4113 CCCTCCAGCCCGCATAGCCGTACCTTCACCTCTGATGCCCGCAGTTGCTGCACCTTGT 4172
 QY 4536 GGACCACTATCAGAGCAGCTGATAGCCCGAGTACATCGGTGGAGTTGTCTTGGACAA 4595
 Db 4173 GGACCACTATCAGAGCAGCTGATAGCCCGAGTACATCGGTGGAGTTGTCTTGGACAA 4232
 QY 4596 ACTGCAATGCAAGTGTCTGAGCCAGGACGATPACAGAGGGTGTCTGGCTGAGAAACACGAG 4655
 Db 4233 ACTGCAATGCAAGTGTCTGAGCCAGGACGATPACAGAGGGTGTCTGGCTGAGAAACACGAG 4292
 QY 4656 GCCCAGCCAGATGCGGAAGCTGTTTCAGCTTGAGCCAGTCTCTGGGACCGGAAGTGAAGA 4715
 Db 4293 GCCCAGCCAGATGCGGAAGCTGTTTCAGCTTGAGCCAGTCTCTGGGACCGGAAGTGAAGA 4352
 QY 4716 TGGACTCTTACCAAGCCCTGAAAGGAGACCATCTCCTCACTTATGGAACCTCTGGAGAA 4775
 Db 4353 TGGACTCTTACCAAGCCCTGAAAGGAGACCATCTCCTCACTTATGGAACCTCTGGAGAA 4412
 QY 4776 GGGCAGCAAAAAGGACCTCTGCCACTCAGCAGCTGAAAGTATCAACAACAGCCCTTGACC 4835
 Db 4413 GGGCAGCAAAAAGGACCTCTGCCACTCAGCAGCTGAAAGTATCAACAACAGCCCTTGACC 4472
 QY 4836 CTTGAGTCTTGGCTTGGCTGACCCCTCTTGGGTCTCAGTTTCTCTCTGCAACCAAG 4895
 Db 4473 CTTGAGTCTTGGCTTGGCTGACCCCTCTTGGGTCTCAGTTTCTCTCTGCAACCAAG 4532
 QY 4896 TTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAACCTTGTATGATGCTTGTCTGG 4955
 Db 4533 TTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGGAACCTTGTATGATGCTTGTCTGG 4592
 QY 4956 GCATTATGTGTCCATGTCAGGGATGCCACAGGGGCCCCAGTCCAGGTGGCCTAACAGCA 5015
 Db 4593 GCATTATGTGTCCATGTCAGGGATGCCACAGGGGCCCCAGTCCAGGTGGCCTAACAGCA 4652
 QY 5016 TCTCAGGGAATGTCCATCTGGAGCTGGCAAGCCCTCGAGATCCATCCAGGCGCAAGAGAAATA 5135
 Db 4653 TCTCAGGGAATGTCCATCTGGAGCTGGCAAGCCCTCGAGATCCATCCAGGCGCAAGAGAAATA 4772
 QY 5136 GGAGGACATGGAACCAATTTGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5195
 Db 4773 GGAGGACATGGAACCAATTTGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4832
 QY 5196 TCAGCGTGGAGGCGCACGTGGATTTCTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5255
 Db 4833 TCAGCGTGGAGGCGCACGTGGATTTCTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4892
 QY 5256 AACAGAGTAAAGTGAAGAGAGTTTATTCAGAAAAATAAGAGAGTATCAAGCTCTTTTAG 5315
 Db 4893 AACAGAGTAAAGTGAAGAGAGTTTATTCAGAAAAATAAGAGAGTATCAAGCTCTTTTAG 4952
 QY 5316 AATTTGTCTAGCAGGCTTTCAGTTTTTACAGAAAAACCCCTATATAATTAATAATTTTTT 5375
 Db 4953 AATTTGTCTAGCAGGCTTTCAGTTTTTACAGAAAAACCCCTATATAATTAATAATTTTTT 5012
 QY 5376 ACTTAAATTAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5435
 Db 5013 ACTTAAATTAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5072
 QY 5436 GTTA 5439
 Db 5073 GTTA 5076

US-09-996-617-5	Db	661	GAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGCACGGTGGTAGGAACG	720
; Sequence 5, Application US/09996617	Qy	1243	CCCCACAGGCGCACACAGCCTACAGCCCCACACACCCATCGGAGCCTTCTGTGAGA	1302
; Patent No. US20020128198A1	Db	721	CCCCACAGGCGCACACAGCCTACAGCCCCACACCCACCCATCGGAGCCTTCTGTGAGA	780
; GENERAL INFORMATION:	Qy	1303	GAGAGCCTCTGTTCCACATGCGCCCTGAAAAATGAGGATTTTAAACCAAAAATTCACACAG	1362
; APPLICANT: Bertin, John	Db	781	GAGAGCCTCTGTTCCACATGCGCCCTGAAAAATGAGGATTTTAAACCAAAAATTCACACAG	840
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED	Qy	1363	CTGCTACTTCTACAAAGACCTCACCCACAGAACCCAGATCCCTCGTCAAGAGAACTCG	1422
; FILE REFERENCE: 07334-340001	Db	841	CTGCTACTTCTACAAAGACCTCACCCACAGAACCCAGATCCCTCGTCAAGAGAACTCG	900
; CURRENT APPLICATION NUMBER: US/09/996,617	Qy	1423	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTTGGCCCA	1482
; PRIOR APPLICATION NUMBER: 09/931,071	Db	901	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTTGGCCCA	960
; PRIOR FILING DATE: 2001-08-15	Qy	1483	GGCCTGGATACCCAAAGAACCTCGCATAGTCACTATGACAGGGGCTGCTGAAATGGGAAG	1542
; PRIOR FILING DATE: 1999-10-27	Db	961	GGCCTGGATACCCAAAGAACCTCGCATAGTCACTATGACAGGGGCTGCTGAAATGGGAAG	1020
; PRIOR APPLICATION NUMBER: 09/428,252	Qy	1543	TCACACCTGGCCAGSCAGGCTGAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCCG	1602
; PRIOR FILING DATE: 1999-06-28	Db	1021	TCACACCTGGCCAGSCAGGCTGAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCCG	1080
; NUMBER OF SEQ ID NOS: 10	Qy	1603	TTCCAGCATGCTTCTACTTCTCAGCTCGACAGAGCTGGCCACAGTCCAAAGTGGTGAAGTCTC	1662
; SOFTWARE: FastSeq for Windows Version 4.0	Db	1081	TTCCAGCATGCTTCTACTTCTCAGCTCGACAGAGCTGGCCACAGTCCAAAGTGGTGAAGTCTC	1140
; SEQ ID NO 5	Qy	1663	GCTGAGCTCATCGGAAAGATGGGACAGCCATCCCGCTCCCATTTAGACAGATCCTGTGCT	1722
; LENGTH: 4287	Db	1141	GCTGAGCTCATCGGAAAGATGGGACAGCCATCCCGCTCCCATTTAGACAGATCCTGTGCT	1200
; TYPE: DNA	Qy	1723	AGCCAGAGCGGCTGCTTTCATCCTCGATGGTGTAGATGACGACGAGATGGGCTTTCGAG	1782
; ORGANISM: Homo sapiens	Db	1201	AGCCAGAGCGGCTGCTTTCATCCTCGATGGTGTAGATGACGACGAGATGGGCTTTCGAG	1260
US-09-996-617-5	Qy	1783	GAGCCGAGTTCTGAGCTCTGTCTGACATGGAGCCAGCCACAGCCGGCGGATGCACCTGCTG	1842
Query Match 78.7%; Score 4287; DB 10; Length 4287;	Db	1261	GAGCCGAGTTCTGAGCTCTGTCTGACATGGAGCCAGCCACAGCCGGCGGATGCACCTGCTG	1320
Best Local Similarity 100.0%; Pred. No. 0;	Qy	1843	GGCAGTTTGTCTGGGAAACATATATCTCCGAGGCATCTTCTCTGATCAGGCTCGGACC	1902
Matches 4287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Db	1321	GGCAGTTTGTCTGGGAAACATATATCTCCGAGGCATCTTCTCTGATCAGGCTCGGACC	1380
Qy 523 ATGGCTGGCGGAGCCTGGGGCGCTGGCCCTGTACTTGGAGTTCCCTGAAGAGGAG 582	Qy	1903	ACAGCTCTGCAGAACCTCATCTTCTTTGGAGCAGGCACTTTGGGTAGAGGCTCTTGGGG	1962
Db 1 ATGGCTGGCGGAGCCTGGGGCGCTGGCCCTGTACTTGGAGTTCCCTGAAGAGGAG 60	Db	1381	ACAGCTCTGCAGAACCTCATCTTCTTTGGAGCAGGCACTTTGGGTAGAGGCTCTTGGGG	1440
Qy 583 CTGAAGGAGTTCCAGCTTCTGCTGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642	Qy	1963	TTTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	2022
Db 61 CTGAAGGAGTTCCAGCTTCTGCTGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120	Db	1441	TTTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	1500
Qy 643 ACACCCGCTCAGCAGAGAAAGACGATGGCATGGAGTGGCTCTGTACTGGTGGCTCAG 702	Qy	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGCTCTTGTGCC	2082
Db 121 ACACCCGCTCAGCAGAGAAAGACGATGGCATGGAGTGGCTCTGTACTGGTGGCTCAG 180	Db	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGCTCTTGTGCC	1560
Qy 703 TATGGGAGCAGGGGCTGGGACCTAGCCCTCCATCTGCGAGCAGATGGGCTCAGG 762	Qy	2083	TGGGTGTCTGGCTGGCTGCACTTCTTCTTTGGAGCAGGCACTTTGGGTAGAGGCTCTTGGG	2142
Db 181 TATGGGAGCAGGGGCTGGGACCTAGCCCTCCATCTAGCCCTCCATCTGGGAGCAGTGGGCTCAGG 240	Db	1561	TGGGTGTCTGGCTGGCTGCACTTGGCTGTATGTCAGCAGATGAAGCGGAGGAAAAATCTC	1620
Qy 763 TCACCTGTGGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 822	Qy	2143	ACACTGACTTCCAGAGCACCACCAACCTCTGTCTACATTAACCTTTGCCCAGGCTCTCCAA	2202
Db 241 TCACCTGTGGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 300	Db	1621	ACACTGACTTCCAGAGCACCACCAACCTCTGTCTACATTAACCTTTGCCCAGGCTCTCCAA	1680
Qy 823 AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACTCCACGCGAGTGTAAATGCCCTGG 882	Qy	2203	GCTCAGCCTTTGGGACCCCGAGCTCAGAGACCTCTGTCTCTCTGGCTGTCTGAGGCACTTGG	2262
Db 301 AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACTCCACGCGAGTGTAAATGCCCTGG 360	Db	1681	GCTCAGCCTTTGGGACCCCGAGCTCAGAGACCTCTGTCTCTCTGGCTGTCTGAGGCACTTGG	1740
Qy 883 ATCCATGAATTCCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGGAGCAGCTG 942	Qy	2263	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGTGTAGATGGGGCATC	2322
Db 361 ATCCATGAATTCCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGGAGCAGCTG 420				
Qy 943 CCTGACACATCTGGAGCCGCTGGAGAGAAATCTCTGGCTCACTCTCTCAACAGCTCTT 1002				
Db 421 CCTGACACATCTGGAGCCGCTGGAGAGAAATCTCTGGCTCACTCTCTCAACAGCTCTT 480				
Qy 1003 CCAAGCTCCCGACCATCAGTCTCCAGCCAGGAGTCAACCAACGCCCCCAATCCACA 1062				
Db 481 CCAAGCTCCCGACCATCAGTCTCCAGCCAGGAGTCAACCAACGCCCCCAATCCACA 540				
Qy 1063 GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGCAGGAG 1122				
Db 541 GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGCAGGAG 600				
Qy 1123 GCTCTCTGGACCAATGGGCTCTGGATGAACGTCAGGAATTTACTACAGAGAATCAGA 1182				
Db 601 GCTCTCTGGACCAATGGGCTCTGGATGAACGTCAGGAATTTACTACAGAGAATCAGA 660				
Qy 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGCGAGCGGTGGTAGGAACG 1242				

Db 1741 CAAAAAAGACCCCTTTTCAGTCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC 1800
Qy 2323 ATCTCCACCTTCTTTGAAGATGGGTATTTCTCAAGACACCCCATCCTCTGAGCTACAGC 2382
Db 1801 ATCTCCACCTTCTTTGAAGATGGGTATTTCTCAAGACACCCCATCCTCTGAGCTACAGC 1860
Qy 2383 TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTTGAGCAATGTCTCTATGCTTTGGAGGATGAG 2442
Db 1861 TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTTGAGCAATGTCTCTATGCTTTGGAGGATGAG 1920
Qy 2443 AAGGGAGAGGTAAACATTTCTAATTCGATCATAGATTTGGAAGACGCTAGAGCATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTCTAATTCGATCATAGATTTGGAAGACGCTAGAGCATAT 1980
Qy 2503 GGAATACATGGCTGTTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGTTAAGTGAT 2562
Db 1981 GGAATACATGGCTGTTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGTTAAGTGAT 2040
Qy 2563 GAGGGGAGAGAGATGGAGAACATTTCTCACTGCCGCTGTCTCAGGGGAGGAACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAACATTTCTCACTGCCGCTGTCTCAGGGGAGGAACCTG 2100
Qy 2623 ATGCACTGGGTCCCGTCCCTGACGTGCTGTGAGCCACATCTCTGAGATCCCTCCAC 2682
Db 2101 ATGCACTGGGTCCCGTCCCTGACGTGCTGTGAGCCACATCTCTGAGATCCCTCCAC 2160
Qy 2683 TGCCTGTAGAGACTCGGAACAAACGTTCTCAGACAGATGATGCCCATTTGGAAGAA 2742
Db 2161 TGCCTGTAGAGACTCGGAACAAACGTTCTCAGACAGATGATGCCCATTTGGAAGAA 2220
Qy 2743 ATGGGCATGTGTAGAAAACAGATGGAGCTCTTCTAGTGTGACATTTCTGCAATTAATTC 2802
Db 2221 ATGGGCATGTGTAGAAAACAGATGGAGCTCTTCTAGTGTGACATTTCTGCAATTAATTC 2280
Qy 2803 AGCCGCCACGTGAAGAGTTTCAGCTGATTTGAGGCGAGCAGACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACGTGAAGAGTTTCAGCTGATTTGAGGCGAGCAGACAGATCAACATGGAGC 2340
Qy 2863 CCCACCATGGTAGTCTGTTTCAAGTGGGTCCAGTTCACAGATGCCATTTGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTAGTCTGTTTCAAGTGGGTCCAGTTCACAGATGCCATTTGGCAGATTTCTC 2400
Qy 2923 TTCTCCGTCTCTCAAGTTCACCAAGAACCTGGAAGAGCTGGACCTAAGTGGAACTCGCTG 2982
Db 2401 TTCTCCGTCTCTCAAGTTCACCAAGAACCTGGAAGAGCTGGACCTAAGTGGAACTCGCTG 2460
Qy 2983 AGCCACTGTGCAAGAGTCTTTGTAAGACCTTGAGACGCGCTGCTGCTCTGAGAG 3042
Db 2461 AGCCACTGTGCAAGAGTCTTTGTAAGACCTTGAGACGCGCTGCTGCTCTGAGAG 2520
Qy 3043 ACCCTGGGTTGGCTGGCTGTGGCTTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGG 3102
Db 2521 ACCCTGGGTTGGCTGGCTGTGGCTTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGG 2580
Qy 3103 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCT 2640
Qy 3163 GGAGCCAAACACTTTTGCAGAGACTGAGACGCGAGCTGAAGCTACAGCGACTGACAG 3222
Db 2641 GGAGCCAAACACTTTTGCAGAGACTGAGACGCGAGCTGAAGCTACAGCGACTGACAG 2700
Qy 3223 CTGGTCAGCTGTGGCTTCAGCTGAGTCTGCTGCCAGGACTGGCTCTGCTTAGTGCC 3282
Db 2701 CTGGTCAGCTGTGGCTTCAGCTGAGTCTGCTGCCAGGACTGGCTCTGCTTAGTGCC 2760
Qy 3283 AGCCCCAGCTGAAGAGCTAGACCTGACAGCAACAACTGGATGACCTTGGCGTGCGA 3342
Db 2761 AGCCCCAGCTGAAGAGCTAGACCTGACAGCAACAACTGGATGACCTTGGCGTGCGA 2820
Qy 3343 CTGCTCTGAGGGGCTCAGGCATCTGCTGCAAACTCATACGCTGGGGCTGGAACAG 3402
Db 2821 CTGCTCTGAGGGGCTCAGGCATCTGCTGCAAACTCATACGCTGGGGCTGGAACAG 2880

Qy 3403 ACAACTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAAACCTCAG 3462
Db 2881 ACAACTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAAACCTCAG 2940
Qy 3463 CTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTAGAGGGCTGATAGG 3522
Db 2941 CTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTAGAGGGCTGATAGG 3000
Qy 3523 GGAGAGATGATTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGG 3582
Db 3001 GGAGAGATGATTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGG 3060
Qy 3583 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGGACGTGAGCAAGATCTTCCCAATT 3642
Db 3061 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGGACGTGAGCAAGATCTTCCCAATT 3120
Qy 3643 GCTGAGATTGAGAGGAAAGCTCCCAAGAGTGTACCGGTGGAACTCTTGTGCGTGCCT 3702
Db 3121 GCTGAGATTGAGAGGAAAGCTCCCAAGAGTGTACCGGTGGAACTCTTGTGCGTGCCT 3180
Qy 3703 TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3762
Db 3181 TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3240
Qy 3763 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3822
Db 3241 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3300
Qy 3823 CACTTCCCTGTAGCTGGCTCTCAAGCTGGGCCAACACGGGTCTCTGTTGTGATGAGA 3882
Db 3301 CACTTCCCTGTAGCTGGCTCTCAAGCTGGGCCAACACGGGTCTCTGTTGTGATGAGA 3360
Qy 3883 GAAGCGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC 3942
Db 3361 GAAGCGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC 3420
Qy 3943 CCACAGCAGCTGATGGTGGCAGGGCTCTCTGTGGACATCAAGGCTGAGCCTGGAGCT 4002
Db 3421 CCACAGCAGCTGATGGTGGCAGGGCTCTCTGTGGACATCAAGGCTGAGCCTGGAGCT 3480
Qy 4003 GTGGAAGCTGTGCACCTCCCTCAGCTTTGTGCTCTCAAGGGGGCCATGTGAGACATCC 4062
Db 3481 GTGGAAGCTGTGCACCTCCCTCAGCTTTGTGCTCTCAAGGGGGCCATGTGAGACATCC 3540
Qy 4063 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAACCCAGCCAGGCTG 4122
Db 3541 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAACCCAGCCAGGCTG 3600
Qy 4123 GAGCTGCATCACATAGTTTCTGGAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 4182
Db 3601 GAGCTGCATCACATAGTTTCTGGAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 3660
Qy 4183 ATGATCCCAATATGCGCTCGCTTCAATCCCGTCACTCTGTGGTGTGCTTTTACACCGC 4242
Db 3661 ATGATCCCAATATGCGCTCGCTTCAATCCCGTCACTCTGTGGTGTGCTTTTACACCGC 3720
Qy 4243 GTCCATCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTACTGCTTCCATTGG 4302
Db 3721 GTCCATCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTACTGCTTCCATTGG 3780
Qy 4303 AAGGAACCTGAGGCTCTGCTATCGAAGCCCTGGAGAACCCAGCTTCTCTCGAGTTCTAC 4362
Db 3781 AAGGAACCTGAGGCTCTGCTATCGAAGCCCTGGAGAACCCAGCTTCTCTCGAGTTCTAC 3840
Qy 4363 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTG 4422
Db 3841 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTG 3900
Qy 4423 GTTGGGAGGCTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCTCCCTCA 4482
Db 3901 GTTGGGAGGCTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCTCCCTCA 3960

QY 4483 GCCCGCATAGCCGTACTCTTCACTCTGGATGCCCGCCGACGTTGCTGACATTTTGTGGACGAG 4542
DB 3961 GCCCGCATAGCCGTACTCTTCACTCTGGATGCCCGCCGACGTTGCTGACATTTTGTGGACGAG 4020
QY 4543 TATCGAGACGCTGATAGCCCGGAGTGACATCGGTGGAGGTTGCTTGGACAACTGCAAT 4602
DB 4021 TATCGAGACGCTGATAGCCCGGAGTGACATCGGTGGAGGTTGCTTGGACAACTGCAAT 4080
QY 4603 GGACAGGTGCTGAGCCGAGGAGCAGTACGAGAGGGTCTGGCTCAGAAACGAGAGGCCGACG 4662
DB 4081 GGACAGGTGCTGAGCCGAGGAGCAGTACGAGAGGGTCTGGCTCAGAAACGAGAGGCCGACG 4140
QY 4663 CAGATCGGAAGCTGTTTACGTTTGAAGCCAGTCTGGGACCGGAAGTGCGAAAGTGAATGCACTC 4722
DB 4141 CAGATCGGAAGCTGTTTACGTTTGAAGCCAGTCTGGGACCGGAAGTGCGAAAGTGAATGCACTC 4200
QY 4723 TACCAAGCCCTGAAGGAGAGCCCATCTCACCTCATTTATGGAATCTGGGAGAGGGCAGC 4782
DB 4201 TACCAAGCCCTGAAGGAGAGCCCATCTCACCTCATTTATGGAATCTGGGAGAGGGCAGC 4260
QY 4783 AAAAAAGGACTCTCTGCCACTCAGCAGC 4809
DB 4261 AAAAAAGGACTCTCTGCCACTCAGCAGC 4287

RESULT 7
US-09-931-071-5
; Sequence 5, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-931-071-5

Query Match 78.7%; Score 4287; DB 10; Length 4287;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 4287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTCTCTGAAGAGGAGGAG 582
DB 1 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTCTCTGAAGAGGAGGAG 60
QY 583 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
DB 61 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAGACGAGTGGCATGGAGGTGCCCTCGTACCTGTGGTCTAG 702
DB 121 ACACCCGCTCAGCCAGAGAGACGAGTGGCATGGAGGTGCCCTCGTACCTGTGGTCTAG 180
QY 703 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGG 762
DB 181 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGG 240
QY 763 TCACTGTGCCGCCAAGCCAGGAGGGGAGGCGACCTCTCCCTCATTTCCCTACAGCCCA 822
DB 241 TCACTGTGCCGCCAAGCCAGGAGGGGAGGCGACCTCTCCCTCATTTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCGAGCCAAACCCACCTTCAACCGAGTGTATTCGCTGG 882

DB 301 AGTGAACCCCACTGGGGTCTCCGAGCCAAACCCACCTCCACCGCAGTGTCTAATGCCCTGG 360
QY 883 ATCCATGAATTCGCGGGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTTCAGACAGCTG 942
DB 361 ATCCATGAATTCGCGGGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTTCAGACAGCTG 420
QY 943 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCCTCCTCTACCAAGCTCTT 1002
DB 421 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCCTCCTCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAGACCATGAGTCTCCAAGCCAGAGTCAACCAAGCGCCGCCACATCCACA 1062
DB 481 CCAAGCTCCCAGACCATGAGTCTCCAAGCCAGAGTCAACCAAGCGCCGCCACATCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCCGACCTAGCACCCAGAGAGCAGAG 1122
DB 541 GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCCGACCTAGCACCCAGAGAGCAGAG 600
QY 1123 GCTCTGGGACCCCAATGGCCTCTGGATGAAACGTCAGGAATTTTACTACACAGAAATCAGA 1182
DB 601 GCTCTGGGACCCCAATGGCCTCTGGATGAAACGTCAGGAATTTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGGACGCGTGGTAGAAG 1242
DB 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGGACGCGTGGTAGAAG 720
QY 1243 CCCCCACAGGGCCACACAGAGCTTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA 1302
DB 721 CCCCCACAGGGCCACACAGAGCTTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA 780
QY 1303 GAGAGCCTCTGTTCCACATGGCCTGGAAATAGAGGATTTTAAACCAAAATTTACACAG 1362
DB 781 GAGAGCCTCTGTTCCACATGGCCTGGAAATAGAGGATTTTAAACCAAAATTTACACAG 840
QY 1363 CTGCTACTTCTACAAGACCTCACCCAGAGCCCAAGATCCCTGCTGTCAGAGAGAGCTGG 1422
DB 841 CTGCTACTTCTACAAGACCTCACCCAGAGCCCAAGATCCCTGCTGTCAGAGAGAGCTGG 900
QY 1423 CTTGATTTATGTGGAGGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTTATTTGCCCA 1482
DB 901 CTTGATTTATGTGGAGGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTTATTTGCCCA 960
QY 1483 GGCCTGGATACCCAAAGAACCTTCGATATGTCATATCTGAGGGGGCTGCTGGAAATTTGGGAAG 1542
DB 961 GGCCTGGATACCCAAAGAACCTTCGATATGTCATATCTGAGGGGGCTGCTGGAAATTTGGGAAG 1020
QY 1543 TCAACACTGGCCAGCAGGTGAAGAACCTGGGGGAGAGCCAGCTGTATGGGACCGC 1602
DB 1021 TCAACACTGGCCAGCAGGTGAAGAACCTGGGGGAGAGCCAGCTGTATGGGACCGC 1080
QY 1603 TTCCAGCATGCTCTTCTACTTTCAGCTGTCAGAGAGCTGGGCCAGTCCCAAGGTGGTGTCTC 1662
DB 1081 TTCCAGCATGCTCTTCTACTTTCAGCTGTCAGAGAGCTGGGCCAGTCCCAAGGTGGTGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCCTGTCT 1722
DB 1141 GCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCCTGTCT 1200
QY 1723 AGGCCAGAGCGGCTGCTCTTTCATCTCGATGCTGTAGATGAGCCAGGATGGGTCTTCGAG 1782
DB 1201 AGGCCAGAGCGGCTGCTCTTTCATCTCGATGCTGTAGATGAGCCAGGATGGGTCTTCGAG 1260
QY 1783 GAGCCGAGTTCGAGCTCTGCTGCACTGGAGCCAGCCACAGCCCGCGGATGCATCTGCTG 1842
DB 1261 GAGCCGAGTTCGAGCTCTGCTGCACTGGAGCCAGCCACAGCCCGCGGATGCATCTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAATATATCTTCCGAGGATCCTTCTCTGATCAGCGCTCGGACC 1902
DB 1321 GGCAGTTTGTGGGAAAATATATCTTCCGAGGATCCTTCTCTGATCAGCGCTCGGACC 1380
QY 1903 ACAGCTCTGCAAGAACTCATCTCTTCTTTGGAGAGGAGGAGGAGTGGGTAGAGTCTCTGGG 1962
DB 1381 ACAGCTCTGCAAGAACTCATCTCTTCTTTGGAGAGGAGGAGGAGTGGGTAGAGTCTCTGGG 1440

Db	3601	GAGCTGCATCACATAGTTCTTGGAAACCCAGAGTCTCCCGCTGGAGTCTCTCTGAAA	3660
Qy	4183	ATGATCCATAATGCTCGCTTCAATCCCGTCACTCTGTGTGTGTTTACACCGC	4242
Db	3661	ATGATCCATAATGCTCGCTTCAATCCCGTCACTCTGTGTGTGTTTACACCGC	3720
Qy	4243	GTCCATCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAATGCTCCATTCGG	4302
Db	3721	GTCCATCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAATGCTCCATTCGG	3780
Qy	4303	AAGGAAGTGGAGTCTGCTATCGAAGCCCTGGAGAAGACAGTGTTCGAGAGTCTAC	4362
Db	3781	AAGGAAGTGGAGTCTGCTATCGAAGCCCTGGAGAAGACAGTGTTCGAGAGTCTAC	3840
Qy	4363	GTTCGGCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGCAAGAAAGTGAATCTG	4422
Db	3841	GTTCGGCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGCAAGAAAGTGAATCTG	3900
Qy	4423	GTTCGGAGGCTTGGTGAACCCAGAGATCTCATGCTGCAACTACTCTGATCCCTCCA	4482
Db	3901	GTTCGGAGGCTTGGTGAACCCAGAGATCTCATGCTGCAACTACTCTGATCCCTCCA	3960
Qy	4483	GCCGATAGCCGATCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT	4542
Db	3961	GCCGATAGCCGATCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT	4020
Qy	4543	TATCGAGAGCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4602
Db	4021	TATCGAGAGCAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4080
Qy	4603	GGACAGGTGTGAGCCAGGAGCAGTACGAGAGGCTGCTGCTGAGAACACGAGGCCAGC	4662
Db	4081	GGACAGGTGTGAGCCAGGAGCAGTACGAGAGGCTGCTGCTGAGAACACGAGGCCAGC	4140
Qy	4663	CAGATCGGAAGCTGTTCAGTGTGAGCCAGTCTGCGGACCGGAAGTGCAGAAATGAGCTC	4722
Db	4141	CAGATCGGAAGCTGTTCAGTGTGAGCCAGTCTGCGGACCGGAAGTGCAGAAATGAGCTC	4200
Qy	4723	TACCAAGCCCTGAAGAGAGCCATCTCACCTCATTTATGGAATCTGCGGAGAGGCGAGC	4782
Db	4201	TACCAAGCCCTGAAGAGAGCCATCTCACCTCATTTATGGAATCTGCGGAGAGGCGAGC	4260
Qy	4783	AAAAAGGACTCTGCACTCAGCAGC	4809
Db	4261	AAAAAGGACTCTGCACTCAGCAGC	4287

RESULT 8
US-09-388-221-1
; Sequence 1, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4422)
US-09-388-221-1

Query Match 76.2%; Score 4148; DB 10; Length 4422;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

Qy	523	ATGGCTGGCGAGCCTGGGGCGGCTGGCTGTCTTACTTGGAGTTCTCTGAAGAGAGGAG	582
Db	1	ATGGCTGGCGAGCCTGGGGCGGCTGGCTGTCTTACTTGGAGTTCTCTGAAGAGAGGAG	60
Qy	583	CTGAAGAGATTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGAGTCTTTCGGGTGAG	642
Db	61	CTGAAGAGATTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGAGTCTTTCGGGTGAG	120
Qy	643	ACACCCGCTCAGCCAGAGAGACAGAGTGGCATGGAGTGGCTCTGTACTCTGTGGCTCAG	702
Db	121	ACACCCGCTCAGCCAGAGAGACAGAGTGGCATGGAGTGGCTCTGTACTCTGTGGCTCAG	180
Qy	703	TATGGGAGACGCGGGCCCTGGGACCTAGCCCTCCATCTCTGGAGCAGATGGGGCTGAGG	762
Db	181	TATGGGAGACGCGGGCCCTGGGACCTAGCCCTCCATCTCTGGAGCAGATGGGGCTGAGG	240
Qy	763	TCACTGTGCGCCCAAGCCAGAGAGGGGAGGCGAGGCACTCTCCCTCATTTCCCTTACAGCCCA	822
Db	241	TCACTGTGCGCCCAAGCCAGAGAGGGGAGGCGAGGCACTCTCCCTCATTTCCCTTACAGCCCA	300
Qy	823	AGTGAACCCCACTTGGGGTCTCCCAAGCCCAACCCACTCCACCGCAGTGTCTAATGCCCTGG	882
Db	301	AGTGAACCCCACTTGGGGTCTCCCAAGCCCAACCCACTCCACCGCAGTGTCTAATGCCCTGG	360
Qy	883	ATCCATGAATTCGCGCGGGGTGACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG	942
Db	361	ATCCATGAATTCGCGCGGGGTGACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG	420
Qy	943	CTGTACATCTGGACGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTT	1002
Db	421	CTGTACATCTGGACGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTT	480
Qy	1003	CCAAAGCTCCCAAGCAGTGTCTCCAAAGCAGGAGTCAACCAAGCCGCCCATCCACA	1062
Db	481	CCAAAGCTCCCAAGCAGTGTCTCCAAAGCAGGAGTCAACCAAGCCGCCCATCCACA	540
Qy	1063	GCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCTAGACCCAGAGAGAGGAG	1122
Db	541	GCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCTAGACCCAGAGAGAGGAG	600
Qy	1123	GCTCTGGGACCCCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACAGAAATCAGA	1182
Db	601	GCTCTGGGACCCCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACAGAAATCAGA	660
Qy	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCGCCCATGGGCGAGCGTGTAGGAAG	1242
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCGCCCATGGGCGAGCGTGTAGGAAG	720
Qy	1243	CCCCCAGAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA	1302
Db	721	CCCCCAGAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA	780
Qy	1303	GAGAGCTCTGTTCACATGGCCCTGGAAATAGGATTTTAAACCAAAATTCACAG	1362
Db	781	GAGAGCTCTGTGTTCACATGGCCCTGGAAATAGGATTTTAAACCAAAATTCACAG	840
Qy	1363	CTGCTACTTCTACAAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG	1422
Db	841	CTGCTACTTCTACAAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG	900
Qy	1423	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTGGCCCA	1482
Db	901	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTGGCCCA	960
Qy	1483	GGCCTGGATACCCAAAGAACCTCGATAGTATCTGACAGGGGCTGCTGGAATTTGGGAG	1542
Db	961	GGCCTGGATACCCAAAGAACCTCGATAGTATCTGACAGGGGCTGCTGGAATTTGGGAG	1020
Qy	1543	TCAACACTGGCCAGGAGTGAAGAGCTCGGGGAGAGGCGAGCTGTATGGGACCGC	1602
Db	1021	TCAACACTGGCCAGGAGTGAAGAGCTCGGGGAGAGGCGAGCTGTATGGGACCGC	1080
Qy	1603	TTCCAGCATGTCTTCTTACTTACGTGCAGAGAGCTGGCCCGAGTCCCAAGGTGTGATCTC	1662

Qy	1183	GAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGGCCCCATGGGCAGCGCGTGGTAGGAACG	1242
Db	661	GAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGGCCCCATGGGCAGCGTGGTAGGAACG	720
Qy	1243	CCCCACAGGGGCAACACAGGCTACAGCGCCACACACACCCATGGGAGCCTTCTGTGAGA	1302
Db	721	CCCCACAGGGGCAACACAGGCTACAGCGCCACACACACCCATGGGAGCCTTCTGTGAGA	780
Qy	1303	GAGAGCCTCTGTTCACACATGSCCTGGAAATAAGAGATTTTAAACAAAATTCACACAG	1362
Db	781	GAGAGCCTCTGTTCACATGSCCTGGAAATAAGAGATTTTAAACAAAATTCACACAG	840
Qy	1363	CTGCTACTTCTACAAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAAAGCTGG	1422
Db	841	CTGCTACTTCTACAAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAAAGCTGG	900
Qy	1423	CTGTATTATGTGGAGGAGATCGAGGACATTTAATTGAGATCGAGACTTATTTGGCCCA	1482
Db	901	CTGTATTATGTGGAGGAGATCGAGGACATTTAATTGAGATCGAGACTTATTTGGCCCA	960
Qy	1483	GGCCTGATACCCAAAGAACCTCGCATAGTCTATCTGCAGGGGGCTGCTGGAATTGGGAAG	1542
Db	961	GGCCTGATACCCAAAGAACCTCGCATAGTCTATCTGCAGGGGGCTGCTGGAATTGGGAAG	1020
Qy	1543	TCAACACTGGCCAGGAGGTGAAGAAAGCCTGGGGGAGAGCCAGCTGTATGGGACCGC	1602
Db	1021	TCAACACTGGCCAGGAGGTGAAGAAAGCCTGGGGGAGAGCCAGCTGTATGGGACCGC	1080
Qy	1603	TTCCAGCATGCTTCTACTTCAAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGAGTCTC	1662
Db	1081	TTCCAGCATGCTTCTACTTCAAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGAGTCTC	1140
Qy	1663	GCTGAGCTCATCGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1722
Db	1141	GCTGAGCTCATCGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1200
Qy	1723	AGGCCAGAGGGCTGCTCTTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTCAG	1782
Db	1201	AGGCCAGAGGGCTGCTCTTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTCAG	1260
Qy	1783	GAGCCGAGTTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCGCGGATGCACTGCTG	1842
Db	1261	GAGCCGAGTTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCGCGGATGCACTGCTG	1320
Qy	1843	GGCAGTTTGTGGGAAAACATACTTCCGAGGCACTCCTCTGTATCAAGGCTCGGACC	1902
Db	1321	GGCAGTTTGTGGGAAAACATACTTCCGAGGCACTCCTCTGTATCAAGGCTCGGACC	1380
Qy	1903	ACAGCTCTGCAGAACCTCATCTCTTTGGAGCAGGACAGCTGGGGTAGAGTCTCTGGGG	1962
Db	1381	ACAGCTCTGCAGAACCTCATCTCTTTGGAGCAGGACAGCTGGGGTAGAGTCTCTGGGG	1440
Qy	1963	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	2022
Db	1441	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	1500
Qy	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTCTTGTGCC	2082
Db	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTCTTGTGCC	1560
Qy	2083	TGGGTGCTCTGGCTGGCTGCTGCTGATGAGCAGATGAAGCGGAAGGAAAATCTC	2142
Db	1561	TGGGTGCTCTGGCTGGCTGCTGCTGATGAGCAGATGAAGCGGAAGGAAAATCTC	1620
Qy	2143	ACACTGACTTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGCCAGGCTCTCCAA	2202
Db	1621	ACACTGACTTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGCCAGGCTCTCCAA	1680
Qy	2203	GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGTCTCTTGGCTGTCTGAGGGCATCTGG	2262
Db	1681	GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGTCTCTTGGCTGTCTGAGGGCATCTGG	1740

QY	2263	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	2322
DB	1741	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	1800
QY	2323	ATCTCCACCTTCCTGAGATGGGTATCTTTCAGAGCACCCCATCCCTCTGAGCTACAGC	2382
DB	1801	ATCTCCACCTTCCTGAGATGGGTATCTTTCAGAGCACCCCATCCCTCTGAGCTACAGC	1860
QY	2383	TTCAATTCACCTCTGTTTTCGAAGAGTCTTTCAGCAATGTCCTATGCTTTGGAGGATGAG	2442
DB	1861	TTCAATTCACCTCTGTTTTCGAAGAGTCTTTCAGCAATGTCCTATGCTTTGGAGGATGAG	1920
QY	2443	AAGGGAGAGGTAAACAATTCATTAATGATCATATAGATTTGGAAAAGACGCTAGAAAGCATAT	2502
DB	1921	AAGGGAGAGGTAAACAATTCATTAATGATCATATAGATTTGGAAAAGACGCTAGAAAGCATAT	1980
QY	2503	GGAAATACATGGCCTGTTTGGGGCATCAACACACGTTTCCCTATTGGGCCCTGTTAAAGTAT	2562
DB	1981	GGAAATACATGGCCTGTTTGGGGCATCAACACACGTTTCCCTATTGGGCCCTGTTAAAGTAT	2040
QY	2563	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGCGTGTCTCAGGSGAGGAACCTTG	2622
DB	2041	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGCGTGTCTCAGGSGAGGAACCTTG	2100
QY	2623	ATGCAGTGGGTCCGTCCTCTGCAGCTGCTGCTGCAGCACACTCTCTGAGTCCCTCCAC	2682
DB	2101	ATGCAGTGGGTCCGTCCTCTGCAGCTGCTGCTGCAGCACACTCTCTGAGTCCCTCCAC	2160
QY	2683	TGCTTTGACAGACTCGGAACAAAAAGTTTCTCTGACACNAGTGATGCGCCATTTTCGAAGAA	2742
DB	2161	TGCTTTGACAGACTCGGAACAAAAAGTTTCTCTGACACNAGTGATGCGCCATTTTCGAAGAA	2220
QY	2743	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATCTTCTGCATTAATAATTC	2802
DB	2221	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATCTTCTGCATTAATAATTC	2280
QY	2803	AGCCGCCACGTGAAGAAGCTTTCAGCTGATTTAGGGGAGGCAGCACAGATCAACATGGAGC	2862
DB	2281	AGCCGCCACGTGAAGAAGCTTTCAGCTGATTTAGGGGAGGCAGCACAGATCAACATGGAGC	2340
QY	2863	CCNACNATGTTAGTCTGTTTCAGGTGGTCCAGTACAGATGCGCTATTGGCAGATTCTC	2922
DB	2341	CCNACNATGTTAGTCTGTTTCAGGTGGTCCAGTACAGATGCGCTATTGGCAGATTCTC	2400
QY	2923	TTCTCCGTCTCTCAAGGTCACACAGAAACCTGAGAGAGCTGGACCTTAAGTGGAAACTCGCTG	2982
DB	2401	TTCTCCGTCTCTCAAGGTCACACAGAAACCTGAGAGAGCTGGACCTTAAGTGGAAACTCGCTG	2460
QY	2983	AGCCACTCTGCAGTGAAGAGTCTTTGTGAAGACCTCTGAGAGCCCTCGCTCCCTCTGGAG	3042
DB	2461	AGCCACTCTGCAGTGAAGAGTCTTTGTGAAGACCTCTGAGAGCCCTCGCTCCCTCTGGAG	2520
QY	3043	ACCTTGGGTTTGGCTGTGTGCGCTCAACGTGTGAGGACTGCAAGAGACCTTGCCCTTTGGG	3102
DB	2521	ACCTTGGGTTTGGCTGTGTGCGCTCAACGTGTGAGGACTGCAAGAGACCTTGCCCTTTGGG	2580
QY	3103	CTGAGAGCCAAACAGACCCCTGACCCGAGCTGGACCTTGAGTCTCAATGTGCTCACGGATGCT	3162
DB	2581	CTGAGAGCCAAACAGACCCCTGACCCGAGCTGGACCTTGAGTCTCAATGTGCTCACGGATGCT	2640
QY	3163	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCAG	3222
DB	2641	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCAG	2700
QY	3223	CTGCTAGCTGTGGCCTCAGCTGTGACTGTGTCAGGAACTTGCGCTCTGTGCTTAGTGCC	3282
DB	2701	CTGCTAGCTGTGGCCTCAGCTGTGACTGTGTCAGGAACTTGCGCTCTGTGCTTAGTGCC	2760
QY	3283	AGCCCCAGCCTGAAGAGCTAGACCTTGACAGAAACAACTGGATGACGTTGCGCTGCGA	3342
DB	2761	AGCCCCAGCCTGAAGAGCTAGACCTTGACAGAAACAACTGGATGACGTTGCGCTGCGA	2820
QY	3343	CTGCTCTGTGAGGGGCTCAGGCATCTCTGCTCGAAACTCATACGCTGGGGCTGGAACCG	3402

Db 2821 |||||CTGCTCTGTGAGGGCTCAGGCATCTGCTGCAAACTCATACGCGCTGG-----2869
QY 3403 ACAACTCTGAGTGATGAGATGAGCGAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCTCAG 3462
Db 2870 -----2869
QY 3463 CTGCTCATCTTACGACAGCGAAACCAAGTGTGATGACCCCTACTAGAGGGCCTGGATACG 3522
Db 2870 -----GGAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG 2910
QY 3523 GGNAGATGAGTAATAGCAATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCG 3582
Db 2911 GGAGATGAGTAATAGCAATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCG 2970
QY 3583 GCTTCCCATCTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAAT 3642
Db 2971 GCTTCCCATCTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAAT 3030
QY 3643 GCTGAGATTGACAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTGTGCGTGCCT 3702
Db 3031 GCTGAGATTGACAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTGTGCGTGCCT 3090
QY 3703 TCTCCTGCTCTCAAGGGACCTGCATACGAAGCCTTTGGGACTGACGATGACTTCTGG 3762
Db 3091 TCTCCTGCTCTCAAGGGACCTGCATACGAAGCCTTTGGGACTGACGATGACTTCTGG 3150
QY 3763 GGCCCCACAGGGGCTCTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3151 GGCCCCACAGGGGCTCTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3210
QY 3823 CACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACACGGGTCTCTGCTTTGTGATGAGA 3882
Db 3211 CACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACACGGGTCTCTGCTTTGTGATGAGA 3270
QY 3883 GAAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGACACAGTTCTCGGTGAGATCAAC 3942
Db 3271 GAAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGACACAGTTCTCGGTGAGATCAAC 3330
QY 3943 CCACACACAGCTGGATGTGGCAGGSCCTCTGCTGAGCATCAAGCTGAGCCTGGAGCT 4002
Db 3331 CCACACACAGCTGGATGTGGCAGGSCCTCTGCTGAGCATCAAGCTGAGCCTGGAGCT 3390
QY 4003 GTGGAGCTGTGCACCTCCCTCACTTTTGGCTCTCCAGGGGGCCATGTGACACATCC 4062
Db 3391 GTGGAGCTGTGCACCTCCCTCACTTTTGGCTCTCCAGGGGGCCATGTGACACATCC 3450
QY 4063 CTGTTCCAAATGGCCCACTTAAAGAGGAGGGATGCTCTGGAGAGCCAGCCAGGGTG 4122
Db 3451 CTGTTCCAAATGGCCCACTTAAAGAGGAGGGATGCTCTGGAGAGCCAGCCAGGGTG 3510
QY 4123 GAGCTGCATCAGATGTTCTGGAAACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAA 4182
Db 3511 GAGCTGCATCAGATGTTCTGGAAACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAA 3570
QY 4183 ATGATCCATAATGCCCTGCGCTTCACTTCCCGTCACTCTGTGTGTGTGTTTACCAACCGC 4242
Db 3571 ATGATCCATAATGCCCTGCGCTTCACTTCCCGTCACTCTGTGTGTGTGTTTACCAACCGC 3630
QY 4243 GTCATCTCTGAGGAAGTCACTTCCACTTCACTGATCCCAAGTCACTGCTCCATTCGG 4302
Db 3631 GTCATCTCTGAGGAAGTCACTTCCACTTCACTGATCCCAAGTCACTGCTCCATTCGG 3690
QY 4303 AAGGAATGGAGTCTGCTATCAAGCCCTGGAGAGACAGCTGTCTCGGAGTTCTAC 4362
Db 3691 AAGGAATGGAGTCTGCTATCAAGCCCTGGAGAGACAGCTGTCTCGGAGTTCTAC 3750
QY 4363 GTTGGCCACTTTGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGATGAGACTCTG 4422
Db 3751 GTTGGCCACTTTGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGATGAGACTCTG 3810
QY 4423 GTGTGGAGGSCCTTGTGTGAACACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCA 4482
|||||

Db 3811 GTGTGGGAGGCTTGTGTGAACACAGGAGATCTATGCTGCTGCAACTACTGATCCCTCA 3870
QY 4483 GCCGCATAGCGGTACCTTACCTCTGGATGCCCGCAGTTGCTGCACTTTGTGGACCAG 4542
Db 3871 GCCGCATAGCGGTACCTTACCTCTGGATGCCCGCAGTTGCTGCACTTTGTGGACCAG 3930
QY 4543 TATCAGAGCAGCTGATAGCCCGAGTGCATCGTGGAGGTTGCTTGGACAACTGCAT 4602
Db 3931 TATCAGAGCAGCTGATAGCCCGAGTGCATCGTGGAGGTTGCTTGGACAACTGCAT 3990
QY 4603 GGACAGGTGCTGAGCCAGGACAGTACGAGAGGTGCTGGCTGAGAACACAGAGGCCAGC 4662
Db 3991 GGACAGGTGCTGAGCCAGGACAGTACGAGAGGTGCTGGCTGAGAACACAGAGGCCAGC 4050
QY 4663 CAGATCGCGAAGCTGTTTTCAGCTTTCAGCCAGTCTCTGGACCGGAAGTGCAGGACTC 4722
Db 4051 CAGATCGCGAAGCTGTTTTCAGCTTTCAGCCAGTCTCTGGACCGGAAGTGCAGGACTC 4110
QY 4723 TACCAAGCCCTGAAGGAGACCCATCTCCTCATTTATGGAATCTTGGAGAGGGCAGC 4782
Db 4111 TACCAAGCCCTGAAGGAGACCCATCTCCTCATTTATGGAATCTTGGAGAGGGCAGC 4170
QY 4783 AAAAAAGGAGCTCCTGCCACTCAGCAGCTGA 4812
Db 4171 AAAAAAGGAGCTCCTGCCACTCAGCAGCTGA 4200

RESULT 10
US-09-388-221-5
; Sequence 5, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4332)
US-09-388-221-5

Query Match 72.7%; Score 3958; DB 10; Length 4332;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 4200; Conservative 0; Mismatches 0; Indels 222; Gaps 2;

QY 523 ATGGCTGGCGGAGCCTGGGGCCGCTGGCTGTTTACTTGGAGTTCTTGAAGAGGAGAG 582
Db 1 ATGGCTGGCGGAGCCTGGGGCCGCTGGCTGTTTACTTGGAGTTCTTGAAGAGGAGAG 60
QY 583 CTGAAGGAGTCCAGCTTCTGCTCCCAATAAGCGCACTCCAGAGCTCTTCGGGTGAG 642
Db 61 CTGAAGGAGTCCAGCTTCTGCTCCCAATAAGCGCACTCCAGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAAAGAGTGGCAGTGGAGTGGCCTCGTACCTGGTGCCTAG 702
Db 121 ACACCCGCTCAGCCAGAGAAAGAGTGGCAGTGGAGTGGCCTCGTACCTGGTGCCTAG 180
QY 703 TATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 762
Db 181 TATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 240
QY 763 TCACCTGTGGCCCAAGCCAGGAGGGCAGGCTCTCCCTCATTTCCCTACAGCCCA 822
Db 241 TCACCTGTGGCCCAAGCCAGGAGGGCAGGCTCTCCCTCATTTCCCTACAGCCCA 300
QY 823 AGTGAACCCACCTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTCTAATGCTGCG 882

Db	301	AGTGAACCCACCTGGGCTCTCCAGCCAAACCCACCTCCAGCGAGTCTAATGCCCTGG	360		1381	ACAGCTCTGCAGAAACCTCAATTCCTTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGG	1440
Qy	883	ATCCATGAATTCGCGGGGGTGCACCCAGGGCTCAGAGAAGGTTTTGAGACAGCTG	942		1963	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA	2022
Db	361	ATCCATGAATTCGCGGGGGTGCACCCAGGGCTCAGAGAAGGTTTTGAGACAGCTG	420		1441	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA	1500
Qy	943	CTTGACACATCTGACGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT	1002		2023	ATTAGAGCCTTTAGGTGGTCAAAATCAAAACAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
Db	421	CTTGACACATCTGACGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT	480		1501	ATTAGAGCCTTTAGGTGGTCAAAATCAAAACAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
Qy	1003	CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA	1062		2083	TGGGTGTCTGGCTGGCCTGCATCTTGCTGATGACAGAGATGAAGCGAAGAAAATCTC	2142
Db	481	CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA	540		1561	TGGGTGTCTGGCCTGCATCTTGCTGATGACAGAGATGAAGCGAAGAAAATCTC	1620
Qy	1063	GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCCAGAGAGCAGGAG	1122		2143	ACACTGACTTCCAAGCACCACCAACCTCTGTCTACATTTACCTTTGCCAGGCTCTCAA	2202
Db	541	GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCCAGAGAGCAGGAG	600		1621	ACACTGACTTCCAAGCACCACCAACCTCTGTCTACATTTACCTTTGCCAGGCTCTCAA	1680
Qy	1123	GCTCTGGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACACAGAAATCAGA	1182		2203	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGGCATCTGG	2262
Db	601	GCTCTGGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACAGAAATCAGA	660		1681	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGGCATCTGG	1740
Qy	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGGCAGCGGTGGTAGGAAG	1242		2263	CAAAAAAGAGCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	2322
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGGCAGCGGTGGTAGGAAG	720		1741	CAAAAAAGAGCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	1800
Qy	1243	CCCCCAGCGGACACAGAGCTACAGCCGACACACCCATGGGAGCTTCTGTGAGA	1302		2323	ATCTCCACTTCTGAAGATGGGTATTTCTCAAGAGCACCACCTCCCTCTGAGCTACAGC	2382
Db	721	CCCCCAGCGGACACAGAGCTACAGCCGACACACCCATGGGAGCTTCTGTGAGA	780		1801	ATCTCCACTTCTGAAGATGGGTATTTCTCAAGAGCACCACCTCCCTCTGAGCTACAGC	1860
Qy	1303	GAGAGCTCTGTTCCACATGGCCCTCGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362		2383	TTCAATTCACCTCTGTTTCCAAAGAGTTCCTTTCAGCAATGTCTCTGTCTGGAGGATGAG	2442
Db	781	GAGAGCTCTGTTCCACATGGCCCTCGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840		1861	TTCAATTCACCTCTGTTTCCAAAGAGTTCCTTTCAGCAATGTCTCTGTGGAGGATGAG	1920
Qy	1363	CTGCTACTTCTCAAAAGACCTCACCCAGAGCAAGATCCCTGTCTCAAGAGAGCTGG	1422		2443	AAGGGAGAGGTAAACATTTCTAATTCATCATAGATTTGGAAAAAGACGCTAGAACATAT	2502
Db	841	CTGCTACTTCTCAAAAGACCTCACCCAGAGCAAGATCCCTGTCTCAAGAGAGCTGG	900		1921	AAGGGAGAGGTAAACATTTCTAATTCATCATAGATTTGGAAAAAGACGCTAGAACATAT	1980
Qy	1423	CTGTATTTGTGAGAGAGAAATCAGAGACATTTAATTTAGATCAGAGACTTATTTGGCCCA	1482		2503	GGAAATACATGGCTGTTTGGGGCATCAACACAGCTTTCCTATTTGGGGCTGTTAAGTAT	2562
Db	901	CTGTATTTGTGAGAGAGAAATCAGAGACATTTAATTTAGATCAGAGACTTATTTGGCCCA	960		1981	GGAAATACATGGCTGTTTGGGGCATCAACACAGCTTTCCTATTTGGGGCTGTTAAGTAT	2040
Qy	1483	GGCTGGATACCAAGAACTCGATAGTCAATCTGAGGGGCTCTCGAATTTGGGAAG	1542		2563	GAGGGGAGAGAGATGGAGAAACATCTTTCACCTCCGGCTGTCTCAGGGGAGGAAACCTG	2622
Db	961	GGCTGGATACCAAGAACTCGATAGTCAATCTGAGGGGCTCTCGAATTTGGGAAG	1020		2041	GAGGGGAGAGAGATGGAGAAACATCTTTCACCTCCGGCTGTCTCAGGGGAGGAAACCTG	2100
Qy	1543	TCAACACTGGCCAGGAGGAGGAGGAGGCTGGGGAGAGGCGCAGCTGTATGGGACCGC	1602		2623	ATGCAAGTGGGTCCCGTCCCTGAGCTGTCTGTGACGACACTCTCTGGAGTCCCTCCAC	2682
Db	1021	TCAACACTGGCCAGGAGGAGGAGGAGGCTGGGGAGAGGCGCAGCTGTATGGGACCGC	1080		2101	ATGCAAGTGGGTCCCGTCCCTGAGCTGTCTGTGACGACACTCTCTGGAGTCCCTCCAC	2160
Qy	1603	TTCCAGACTGTCTTACTTTAGCTGCAGAGAGCTGGCCAGTCCAGGTGGTGTAGTCTC	1662		2683	TGCTTGTACGAGACTCGGAACAAACAGTTCCTGTGACAAAGTGTGGGCCATTTTCAAGAA	2742
Db	1081	TTCCAGACTGTCTTACTTTAGCTGCAGAGAGCTGGCCAGTCCAGGTGGTGTAGTCTC	1140		2161	TGCTTGTACGAGACTCGGAACAAACAGTTCCTGTGACAAAGTGTGGGCCATTTTCAAGAA	2220
Qy	1663	GCTGAGCTATCGGAAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1722		2743	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATCTTCTGCATTAATTC	2802
Db	1141	GCTGAGCTATCGGAAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1200		2221	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATCTTCTGCATTAATTC	2280
Qy	1723	AGGCCAGAGCGGTCTCTTCTCTCGATGGTGTAGATGAGCCAGATGGGTCTTGCGAG	1782		2803	AGCCGCCAGTGAAGAGCTTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGGAGC	2862
Db	1201	AGGCCAGAGCGGTCTCTTCTCTCGATGGTGTAGATGAGCCAGATGGGTCTTGCGAG	1260		2281	AGCCGCCAGTGAAGAGCTTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGGAGC	2340
Qy	1783	GAGCCGAGTTCTGAGCTGTCTGTGCACTGGAGCAGCCACAGCCGCGGATGCACTGCTG	1842		2863	CCCACCATGGTAGTCTCTTTCAGGTGGGTCCAGTTCAGATGCCATTTGGCAGATTTCTC	2922
Db	1261	GAGCCGAGTTCTGAGCTGTCTGTGCACTGGAGCAGCCACAGCCGCGGATGCACTGCTG	1320		2341	CCCACCATGGTAGTCTCTTTCAGGTGGGTCCAGTTCAGATGCCATTTGGCAGATTTCTC	2400
Qy	1843	GGCAGTTTCTGGGAAACTACTACTCCGAGGCACTCTCTGATCAGGCTCGGACC	1902		2923	TTCTCCGTCTCAAGGTCAACAGAAAACCTGAAGAGCTGGACCTTAAGTGGAAATCTCGGTG	2982
Db	1321	GGCAGTTTCTGGGAAACTACTACTCCGAGGCACTCTCTGATCAGGCTCGGACC	1380		2401	TTCTCCGTCTCAAGGTCAACAGAAAACCTGAAGAGCTGGACCTTAAGTGGAAATCTCGGTG	2460
Qy	1903	ACAGCTCTGCAGAAACCTCAATTCCTTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGG	1962		2983	AGCCACTCTGAGTGAAGAGTCTTTTGAAGACCTTGAGACGCGCTCTCGTGCCTCTCGAG	3042
					2461	AGCCACTCTGAGTGAAGAGTCTTTTGAAGACCTTGAGACGCGCTCTCGTGCCTCTCGAG	2520

QY 3043 ACCCTCGGTTGGCTGGCTGGCTCAGAGTGAAGACCTTGCTTTGGG 3102
 Db 2521 ACCCTCGGTTGGCTGGCTGGCTCAGAGTGAAGACCTTGCTTTGGG 2580
 QY 3103 CTGAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCAGGATGCT 3162
 Db 2581 CTGAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCAGGATGCT 2640
 QY 3163 GGAGCCAAACACCTTTGCCAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCGAG 3222
 Db 2641 GGAGCCAAACACCTTTGCCAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCGAG 2700
 QY 3223 CTGGTCAGTGTGGCTCAGCTGTGATGCTGTGCGAGGACCTGCGCTCTGTGCTTAGTGCC 3282
 Db 2701 CTGGTCAGTGTGGCTCAGCTGTGATGCTGTGCGAGGACCTGCGCTCTGTGCTTAGTGCC 2760
 QY 3283 AGCCCGAGCTGAGAGCTGAGCTGAGCAGCAACCTGGATGACCTGGCGTGGCA 3342
 Db 2761 AGCCCGAGCTGAGAGCTGAGCTGAGCAGCAACCTGGATGACCTGGCGTGGCA 2820
 QY 3343 CTGCTCTGTGAGGGGCTCAGGATCTGCTGCAAACTCATACGCTGGGGCTGGACCAAG 3402
 Db 2821 CTGCTCTGTGAGGGGCTCAGGATCTGCTGCAAACTCATACGCTGG 2869
 QY 3403 ACACTCTGAGTGATGAGTGAGGCAAGAACTGAGGGCCCTGGAGCAGGAGAACTCAG 3462
 Db 2870 2869
 QY 3463 CTGCTCATCTTCCAGCAGCGAACAAGTGTGATGACCCCTACTGAGGGCTGGATAG 3522
 Db 2870 2870
 QY 3523 GGAGAGTGAATAGCATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 3582
 Db 2911 GGAGAGTGAATAGCATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 2970
 QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCGTGAGCAAGATCTTCCCAATT 3642
 Db 2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCGTGAGCAAGATCTTCCCAATT 3030
 QY 3643 GCTGAGATGAGAGAGAGCTCCAGAGGTAGTACCGGTGAACTCTTGTGCGTGCCT 3702
 Db 3031 GCTGAGATGAGAGAGAGCTCCAGAGGTAGTACCGGTGAACTCTTGTGCGTGCCT 3090
 QY 3703 TCTCTGCTCTCAAGGGGACCTGATACGAGGCTTTGGGACTCAGCATGACTTCTG 3762
 Db 3091 TCTCTGCTCTCAAGGGGACCTGATACGAGGCTTTGGGACTCAGCATGACTTCTG 3150
 QY 3763 GGCCCGCAGCGGCTCTGCTGCTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
 Db 3151 GGCCCGCAGCGGCTCTGCTGCTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3210
 QY 3823 CACTTCCCTGTAGTGGCTCTACCGGTGCGCCAAACAGGGTCTCTGCTTTGTGATGAGA 3882
 Db 3211 CACTTCCCTGTAGTGGCTCTACCGGTGCGCCAAACAGGGTCTCTGCTTTGTGATGAGA 3270
 QY 3883 GAAGCGGTGACCTGAGATGAACTTCTGTGCTGGGACAGTTCTGGGTGAGATCAAC 3942
 Db 3271 GAAGCGGTGACCTGAGATGAACTTCTGTGCTGGGACAGTTCTGGGTGAGATCAAC 3330
 QY 3943 CCACAGCAGCTGAGTGGTGGCAGGGCTCTGCTGAGCATCAAGGCTGAGCCTGGAGCT 4002
 Db 3331 CCACAGCAGCTGAGTGGTGGCAGGGCTCTGCTGAGCATCAAGGCTGAGCCTGGAGCT 3390
 QY 4003 GTGGAAGTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGAACAATCC 4062
 Db 3391 GTGGAAGTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGAACAATCC 3450
 QY 4063 CTGTTCCAAATGCCCCACTTAAAGAGGAGGGGATGCTCTCTGAGAGCCAGCGAGGTG 4122
 Db 3451 CTGTTCCAAATGCCCCACTTAAAGAGGAGGGGATGCTCTCTGAGAGCCAGCGAGGTG 3510

QY 4123 GAGCTGCATCATATAGTTCTTGAAAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAA 4182
 Db 3511 GAGCTGCATCATATAGTTCTTGAAAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAA 3570
 QY 4183 ATGATCCATAATGCTGCTTCAATCCCGTCACTCTGTGCTGCTGCTTTTACCAACGC 4242
 Db 3571 ATGATCCATAATGCTGCTTCAATCCCGTCACTCTGTGCTGCTGCTTTTACCAACGC 3630
 QY 4243 GTCCATCTCTGAGAAAGTCACTTCCACCTCTACCTGATCCCAAGTCACTGCTTCCATCGG 4302
 Db 3631 GTCCATCTCTGAGAAAGTCACTTCCACCTCTACCTGATCCCAAGTCACTGCTTCCATCGG 3690
 QY 4303 4302
 Db 3691 AAGGCCATAGATGATCTAGAAATGAAATTCAGTTTGTGCGAATCCCAAGCCACCCCG 3750
 QY 4303 4302
 Db 3751 CTGACCCCACTTATATATGGGCTGTCTTACACTGTGTCTGGGTCTGGTTCAGGATGCTG 3810
 QY 4303 4303
 Db 3811 GAAATACTCCCAAGAACTGGAGCTCTGCTATCGAAGCCCTGGAGAACAGCAGCTGTT 3870
 QY 4351 TCGGAGTTCTAGCTTGGCCACTTTGGGATCAGGATCAGGCTCAAGTGAAGACAGAAA 4410
 Db 3871 TCGGAGTTCTAGCTTGGCCACTTTGGGATCAGGATCAGGCTCAAGTGAAGACAGAAA 3930
 QY 4411 GATGAGACTCTGGTGGAGGCTTGGTGAACAGGAGATCTCATGCTCCCACTACT 4470
 Db 3931 GATGAGACTCTGGTGGAGGCTTGGTGAACAGGAGATCTCATGCTCCCACTACT 3990
 QY 4471 CTGATCCCTCCAGCCGATAGCCGTACCTTCACTCTGATGCTCCCGCAGTTGCTGCAC 4530
 Db 3991 CTGATCCCTCCAGCCGATAGCCGTACCTTCACTCTGATGCTCCCGCAGTTGCTGCAC 4050
 QY 4531 TTTGTGGACCACTGAGAGCAGCTGATAGCCCGAGTACATCGGTGAGGTGTTGTTG 4590
 Db 4051 TTTGTGGACCACTGAGAGCAGCTGATAGCCCGAGTACATCGGTGAGGTGTTGTTG 4110
 QY 4591 GACAACTGCATGAGCAGTGTGAGCCAGGAGCAGTACGAGGGTCTGCTGAGAAC 4650
 Db 4111 GACAACTGCATGAGCAGTGTGAGCCAGGAGCAGTACGAGGGTCTGCTGAGAAC 4170
 QY 4651 ACAGGCCCCAGCCAGATGCGAAGCTGTTTCACTTGAAGCCAGTCTGAGAGGGTCTGCTGAGAAC 4710
 Db 4171 ACAGGCCCCAGCCAGATGCGAAGCTGTTTCACTTGAAGCCAGTCTGAGAGGGTCTGCTGAGAAC 4230
 QY 4711 AAAGATGAGCTCTCAAGCCCTGAAGGAGACCCATCTCCTCATTTATGAACTCTG 4770
 Db 4231 AAAGATGAGCTCTCAAGCCCTGAAGGAGACCCATCTCCTCATTTATGAACTCTG 4290
 QY 4771 GAGAGGGCAGCAAAAAGGACTCTCTGCCACTCAGCAGCTGA 4812
 Db 4291 GAGAGGGCAGCAAAAAGGACTCTCTGCCACTCAGCAGCTGA 4332

RESULT 11

US-09-388-221-9
 ; Sequence 9, Application US/09388221A
 ; Publication NO. US20020192643A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regulat
 ; FILE REFERENCE: P-LJ 3650
 ; CURRENT APPLICATION NUMBER: US/09/388,221A
 ; CURRENT FILING DATE: 1999-09-01
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 4556
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence

; FEATURE: ; NAME/KEY: CDS ; LOCATION: (1) (4362) ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic ; OTHER INFORMATION: Construct ; US-09-388-221-9									
Query Match 62.5%; Score 3400.4; DB 10; Length 4556; Best Local Similarity 93.7%; Pred. No. 0; Mismatches 236; Indels 3; Gaps 1; Matches 3555; Conservative 0									
QY	523	ATGGCTGGGGAGCCCTGGGGCCGCTGGCTGTACTTGGAGTTCCTGAAGAAGGAGGAG	582						
DB	1	ATGGCTGGCGAGCCTGGGGCCGCTGGCTGTACTTGGAGTTCCTGAAGAAGGAGGAG	60						
QY	583	CTGAAGGAGTTCAGCTTCTGTCTGCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	642						
DB	61	CTGAAGGAGTTCAGCTTCTGTCTGCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	120						
QY	643	ACACCCGCTCAGCAGAGAAGAGGAGTGGAGTGGCTCGTACCTGGTGGCTCAG	702						
DB	121	ACACCCGCTCAGCAGAGAAGAGGAGTGGAGTGGCTCGTACCTGGTGGCTCAG	180						
QY	703	TATGGGAGCAGCGGGCCCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG	762						
DB	181	TATGGGAGCAGCGGGCCCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG	240						
QY	763	TCATCTGTGGCCCAAGCCAGGAAGGGGAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA	822						
DB	241	TCATCTGTGGCCCAAGCCAGGAAGGGGAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA	300						
QY	823	AGTGAACCCACCTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTGTAAATGCCCTGG	882						
DB	301	AGTGAACCCACCTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTGTAAATGCCCTGG	360						
QY	883	ATCCATGAATTTGGCGGGGGTGCACCCAGGGCTCAGAGAGAAGGGTTTGGAGACAGCTG	942						
DB	361	ATCCATGAATTTGGCGGGGGTGCACCCAGGGCTCAGAGAGAAGGGTTTGGAGACAGCTG	420						
QY	943	CTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGGCTCACTCTCTACCAAGCTCTT	1002						
DB	421	CTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGGCTCACTCTCTACCAAGCTCTT	480						
QY	1003	CCAAGCTCCCGAGACCATGATCTCAAGCGAGGTCAACCAAGCCGCCCATCCCA	1062						
DB	481	CCAAGCTCCCGAGACCATGATCTCAAGCGAGGTCAACCAAGCCGCCCATCCCA	540						
QY	1063	GCAGTCTGGGAGCTGGGGATCCCACTCTCAGCCAGCTAGCAACCCAGAGAGCAGGAG	1122						
DB	541	GCAGTCTGGGAGCTGGGGATCCCACTCTCAGCCAGCTAGCAACCCAGAGAGCAGGAG	600						
QY	1123	GTCCTGGGACCAATGGCTCTGGATGAAACGTCAAGAAATTTACTACAGAAATCAGA	1182						
DB	601	GTCCTGGGACCAATGGCTCTGGATGAAACGTCAAGAAATTTACTACAGAAATCAGA	660						
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAGCGAGCCCGCCATGGCGAGTGTAGGAACG	1242						
DB	661	GAAAGAGAGAGAGAAATCAGAGAAGCGAGCCCGCCATGGCGAGTGTAGGAACG	720						
QY	1243	CCCCACAGCGCACACAGCTTACAGCCCGCCACCAACCCATGGGAGCTTCTGTGAGA	1302						
DB	721	CCCCACAGCGCACACAGCTTACAGCCCGCCACCAACCCATGGGAGCTTCTGTGAGA	780						
QY	1303	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362						
DB	781	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840						
QY	1363	CTGCTACTTCTCAAGAAGCTCAACCCAGAGCCAGATCCCTGGTCAAGAGAGCTGG	1422						
DB	841	CTGCTACTTCTCAAGAAGCTCAACCCAGAGCCAGATCCCTGGTCAAGAGAGCTGG	900						
QY	1423	CCTGATTTATGTGGAGAGAAATCAGGACATTTAAATTTAGATCAGAGACTTTATTTGGCCCA	1482						

Best Local Similarity 91.3%; Pred. No. 0; Matches 3465; Conservative 0; Mismatches 236; Indels 93; Gaps 2;

QY	523	ATGGCTGGCGGAGCTGGGCGCGCTGTTACTTTGGAGTTCCTGAAGAGGAGGAG	582
DB	1	ATGGCTGGCGGAGCTGGGCGCGCTGTTACTTTGGAGTTCCTGAAGAGGAGGAG	60
QY	583	CTGAAGGAGTTCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTTTCGGGTGAG	642
DB	61	CTGAAGGAGTTCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTTTCGGGTGAG	120
QY	643	ACACCGCTCAGCAGAGAAGCAGATGGAGTGGAGTTCCTGTAACCTGTTGGCTCAG	702
DB	121	ACACCGCTCAGCAGAGAAGCAGATGGAGTGGAGTTCCTGTAACCTGTTGGCTCAG	180
QY	703	TATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG	762
DB	181	TATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG	240
QY	763	TCACTGTGGCGCCCAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	822
DB	241	TCACTGTGGCGCCCAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	300
QY	823	AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACTCCACCGCAGTGTAAATGCCCTGG	882
DB	301	AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACTCCACCGCAGTGTAAATGCCCTGG	360
QY	883	ATCCATGAATTGGCGGGGGTGCACCCAGGGCTCAGAGAAGGTTTGGAGACAGCTG	942
DB	361	ATCCATGAATTGGCGGGGGTGCACCCAGGGCTCAGAGAAGGTTTGGAGACAGCTG	420
QY	943	CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT	1002
DB	421	CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT	480
QY	1003	CCAAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA	1062
DB	481	CCAAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA	540
QY	1063	GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCACCCAGAGAGCGAGGAG	1122
DB	541	GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCACCCAGAGAGCGAGGAG	600
QY	1123	GCTCTGGGACCAATGGCTCTGGATGAACCTCAGGAAATTTACTACACAGAAATCAGA	1182
DB	601	GCTCTGGGACCAATGGCTCTGGATGAACCTCAGGAAATTTACTACACAGAAATCAGA	660
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAGGCGAGGCCCCCATGGGCGAGCGTGGTAGGAACG	1242
DB	661	GAAAGAGAGAGAGAAATCAGAGAAGGCGAGGCCCCCATGGGCGAGCGTGGTAGGAACG	720
QY	1243	CCCCCAGCGGACACAGCTCAGAGCCGACACCCAGCCATGGAGGCTTCTGTGAGA	1302
DB	721	CCCCCAGCGGACACAGCTCAGAGCCGACACCCAGCCATGGAGGCTTCTGTGAGA	780
QY	1303	GAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362
DB	781	GAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840
QY	1363	CTGCTACTTCTCAAAAGACCTACCCCGAGAAGCCCAAGATCCCTGCTCAAGAGAAGCTGG	1422
DB	841	CTGCTACTTCTCAAAAGACCTACCCCGAGAAGCCCAAGATCCCTGCTCAAGAGAAGCTGG	900
QY	1423	CCTGATATGTGGAGAGAAATCAGAGACATTTAAATTTAGATCAGAGCTTATTTGGCCCA	1482
DB	901	CCTGATATGTGGAGAGAAATCAGAGACATTTAAATTTAGATCAGAGCTTATTTGGCCCA	960
QY	1483	GGCTGGATACCAAGAACTCGCATAGTCACTCAGGGGCTCTCTGAAATTTGGGAAG	1542
DB	961	GGCTGGATACCAAGAACTCGCATAGTCACTCAGGGGCTCTCTGAAATTTGGGAAG	1020
QY	1543	TCAACACTGGCCAGGAGTGAAGAGCCTGGGGAGAGGCGAGCTGTATGGGAGCCGC	1602
DB			

DB	1021	TCAACACTGGCCAGCAGGTTGAAGGAAGCCTGGGGAGAGGCGACGCTGTATGGGAGCCGC	1080
QY	1603	TTCCAGCAGTGTCTTCTACTTTCAGCTGCAGAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTC	1662
DB	1081	TTCCAGCAGTGTCTTCTACTTTCAGCTGCAGAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTC	1140
QY	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGGGTCCCATTTAGACAGATCTGTCT	1722
DB	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGGGTCCCATTTAGACAGATCTGTCT	1200
QY	1723	AGGCCAGAGCGGCTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGAG	1782
DB	1201	AGGCCAGAGCGGCTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGAG	1260
QY	1783	GAGCGGAGTTCGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCATCTGCTG	1842
DB	1261	GAGCGGAGTTCGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCATCTGCTG	1320
QY	1843	GGCAGTTTGTGGGAAAAATATATCTCCGAGGCACTCTTCTCGATCAAGGCTCGGACC	1902
DB	1321	GGCAGTTTGTGGGAAAAATATATCTCCGAGGCACTCTTCTCGATCAAGGCTCGGACC	1380
QY	1903	ACAGCTCTGCAGAACCTCATTTCTTGGAGCAGGCAAGCTTGGGTAGAGTCTCTGGG	1962
DB	1381	ACAGCTCTGCAGAACCTCATTTCTTGGAGCAGGCAAGCTTGGGTAGAGTCTCTGGG	1440
QY	1963	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	2022
DB	1441	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	1500
QY	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
DB	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
QY	2083	TGGGTGTCTGGCTGGCTGCACTTGTGCTGATGAGCAGATGAAGCGGAGGAAACTC	2142
DB	1561	TGGGTGTCTGGCTGGCTGCACTTGTGCTGATGAGCAGATGAAGCGGAGGAAACTC	1620
QY	2143	ACACTGACTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAA	2202
DB	1621	ACACTGACTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAA	1680
QY	2203	GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGTGTGAGGGCATCTGG	2262
DB	1681	GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGTGTGAGGGCATCTGG	1740
QY	2263	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGGTAGATGGGCCATC	2322
DB	1741	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGGTAGATGGGCCATC	1800
QY	2323	ATCTCCACCTTTGAAGATGGGTATTTCTCAAGAGACCCCATCTCTGAGCTACAGC	2382
DB	1801	ATCTCCACCTTTGAAGATGGGTATTTCTCAAGAGACCCCATCTCTGAGCTACAGC	1860
QY	2383	TTCAATTACCTCTGTTTCCAAAGAGTCTTTTGAGCAATGTCTCTGTTGAGGATGAG	2442
DB	1861	TTCAATTACCTCTGTTTCCAAAGAGTCTTTTGAGCAATGTCTCTGTTGAGGATGAG	1920
QY	2443	AAGGGAGAGGTAAACATTTCTAAATTTGATCATAGATTTGGAAAGACGCTAGAGCATAT	2502
DB	1921	AAGGGAGAGGTAAACATTTCTAAATTTGATCATAGATTTGGAAAGACGCTAGAGCATAT	1980
QY	2503	GGAAATACATGGCTGTTGGGGCATCAACACAGCTTCTTATTTGGGCTGTGTTAGTGTAT	2562
DB	1981	GGAAATACATGGCTGTTGGGGCATCAACACAGCTTCTTATTTGGGCTGTGTTAGTGTAT	2040
QY	2563	GAGGGGAGAGAGATGGGAAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACTCTG	2622
DB	2041	GAGGGGAGAGAGATGGGAAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACTCTG	2100
QY	2623	ATGCAAGTGGGCTCCCTGCTGAGCTGTGTGTGAGCAGCACTCTCTGAGGTCCCTCCAC	2682
DB	2101	ATGCAAGTGGGCTCCCTGCTGAGCTGTGTGTGAGCAGCACTCTCTGAGGTCCCTCCAC	2160

Db 9 GAAGAAGCTTACGTGATTAGGGCAGGAGCAGACAGATCAATGAGAGCCCAACATGCT 68
QY 2874 AGTCCTGTTTCAGGTGGGTCCAGTACAGATGCTATTGGCAGATTTCTTCTCCGTCTCT 2933
Db AGTCTTGTTCAGGTGGGTCCAGTACAGATGCTATTGGCAGATTTCTTCTCCGTCTCT 128
QY 2934 CAAGGTACACAGAAACCTGAAGAGAGCTGGAACCTTAAGTGGAAACTCGCTGAGCCACTCTGC 2993
Db 129 CAAGGTACACAGAAACCTGAAGAGAGCTGGAACCTTAAGTGGAAACTCGCTGAGCCACTCTGC 188
QY 2994 AGTGAAGAGTCTTTGTAAGACCTGAGACGCTGAGACGCTCGCTGCTTGGGCTGAGAGCCAA 3053
Db 189 AGTGAAGAGTCTTTGTAAGACCTGAGACGCTGAGACGCTCGCTGCTTGGGCTGAGAGCCAA 248
QY 3054 GCGTGGCTGTGGCTTCAAGCTGAGACCTGCAAGGACCTTGGCTTGGGCTGAGAGCCAA 3113
Db 249 GCGTGGCTGTGGCTTCAAGCTGAGACCTGCAAGGACCTTGGCTTGGGCTGAGAGCCAA 308
QY 3114 CCAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGGATGCTGGAGCCAAACA 3173
Db 309 CCAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGGATGCTGGAGCCAAACA 368
QY 3174 CTTTGGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCGACTGGTCAAGCTG 3233
Db 369 CTTTGGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCGACTGGTCAAGCTG 428
QY 3234 TGGCCTCAGCTGACTGCTGCGAGACCTGGCCTCTGTGCTTAGTGCCAGCCCGACGCT 3293
Db 429 TGGCCTCAGCTGACTGCTGCGAGACCTGGCCTCTGTGCTTAGTGCCAGCCCGACGCT 488
QY 3294 GAAGGAGCTAGACCTGACGAGAAACCTGGATGATGCTTGGCGTGGGACTGCTCTGTGA 3353
Db 489 GAAGGAGCTAGACCTGACGAGAAACCTGGATGATGCTTGGCGTGGGACTGCTCTGTGA 548
QY 3354 GGGCTCAGGCTATCTGCTGCTGCAACTATAGCTTGGGGCTGGACAGAACTCTGTAG 3413
Db 549 GGGGCTCA - GCATCTCTGCTGCAAACTATAGCTTGGGGCTGGACAGAACTCTGTAG 607
QY 3414 TGATGAGATGAGCAGGAACCTGAGGCGCTGGAGCAGGAGAAACCTCAGCTCTCATCTT 3473
Db 608 TGATGAGATGAGCAGGAACCTGAGGCGCTGGAGCAGGAGAAACCTCAGCTCTCATCTT 667
QY 3474 CAGCAGCGGAACCAAGTGTGATGACCTCTACTGAGGCGCTGGATACGGGAGAGATGAG 3533
Db 668 CAGCAGCGGAACCAAGTGTGATGACCTCTACTGA - GGCCTGGATACGGGAGAGATGAG 726
QY 3534 TAAATGACATCTCTCACTCAAGCGGAGAGACTCGGATCAGAGAGCGGCTTCCATGT 3593
Db 727 TAAATGACATCTCTCACTCAAGCGGAGAGACTCGGATCAGAGAGCGGCTTCCATGT 786
QY 3594 TGCTCAGGCTAATCTCAAACTCTGACGCTGAGCAGATCTTCCAAATGCTGAGATTGC 3653
Db 787 TGCTCAGGCTAATCTCAAACTCTGACGCTGAGCAGATCTTCCCAATTTGCTGAGATTGC 846
QY 3654 AGAGGAAAGCTCCCAAGAGTAGTACCGGTGGAACCTTGTGCGTGCCTTCTCCTGCCTC 3713
Db 847 AGAGGAAAGCTCCCAAGAGTAGTACCGGTGGAACCTTGTGCGATGCTTCTCCTGCCTC 906
QY 3714 TCAAGGGGACCTGATACAGAGCCTTTTGGGAGCTGACGATGACTTCTGGGGCCCAACGGG 3773
Db 907 TCAAGGGGACCTGATACAGAGCCTTTTGGGAGCTGACGATGACTTCTGGGGCCCAACGGG 966
QY 3774 GCGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTGT 3833
Db 967 GCGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTGT 1026
QY 3834 AGTGGCTCTTACCGCTGCCCCAACACAGCGGTCTCTGCTTTGTGATGAGAGAGCGGTGAC 3893
Db 1027 AGTGGCTCTTACCGCTGCCCCAACACAGCGGTCTCTGCTTTGTGATGAGAGAGCGGTGAC 1086
QY 3894 CGTTGAGATTGAATCTGTGTGGGACCAAGTTCTGTGGGTGAGTCAACCCACAGACAG 3953
Db 1087 CGTTGAGATTGAATCTGTGTGGGACCAAGTTCTGTGGGTGAGTCAACCCACAGACAG 1146

QY 3954 CTGGATGTGTCAGGGCCCTCTCTGGACATCAAGGCTGAGCCTGAGCTGTGGAAGCTGT 4013
Db 1147 CTGGATGTGTCAGGGCCCTCTCTGGACATCAAGGCTGAGCCTGAGCTGTGGAAGCTGT 1206
QY 4014 GCACCTCCCTCACTTTTGGGCTCTCCAAAGGGGGCATTGTGACACATCCCTGTTCCAAAT 4073
Db 1207 GCACCTCCCTCACTTTTGGGCTCTCCAAAGGGGGCATTGTGACACATCCCTGTTCCAAAT 1266
QY 4074 GGGCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGAGCCAGCGGTGGAGCTGCATCA 4133
Db 1267 GGGCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGAGCCAGCGGTGGAGCTGCATCA 1326
QY 4134 CATAGTTCTGGAAGAACCCAGCTTCTCCCCCTTGGGAGTCTCTCTGAAATGATCCATAA 4193
Db 1327 CATAGTTCTGGAAGAACCCAGCTTCTCCCCCTTGGGAGTCTCTCTGAAATGATCCATAA 1386
QY 4194 TGCCCTGGCTTCATTTCCCGTCACTCTGTGGTGTGTTTACACCGCGTCCATCTCTGA 4253
Db 1387 TGCCCTGGCTTCATTTCCCGTCACTCTGTGGTGTGTTTACACCGCGTCCATCTCTGA 1446
QY 4254 GGAAGTCACTTCCACCTCTACCTGATCCCAAGTGACTCTCCATTCCGAAAGGAACTGGA 4313
Db 1447 GGAAGTCACTTCCACCTCTACCTGATCCCAAGTGACTCTCCATTCCGAAAGGAACTGGA 1506
QY 4314 GCTCTGCTATCGAAGCCCTGGAGAGAGCAGCTGTCTCTCGAGTTTCTACGTTGGCCACTT 4373
Db 1507 GCTCTGCTATCGAAGCCCTGGAGAGAGCAGCTGTCTCTCGAGTTTCTACGTTGGCCACTT 1566
QY 4374 GGGATCAGGGATCAGGGTGCAGGTGAAAGACAAGAAAGATGAGACTCTGGTGTGGAGGC 4433
Db 1567 GGGATCAGGGATCAGGGTGCAGGTGAAAGACAAGAAAGATGAGACTCTGGTGTGGAGGC 1626
QY 4434 CTTGGTGAACACAGAGATCTCATGCGCTGCAACTCTGATCCCTCCAGCCCGCATAGC 4493
Db 1627 CTTGGTGAACACAGAGATCTCATGCGCTGCAACTCTGATCCCTCCAGCCCGCATATC 1686
QY 4494 CGTACTTTCAGCTCTGGATGCCCGCAGTTGCTGCATTTTGTGGACCAAGTATCGAGAGCA 4553
Db 1687 CGTACTTTCAGCTCTGGATGCCCGCAGTTGCTGCATTTTGTGGACCAAGTATCGAGAGCA 1746
QY 4554 GCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGGAACAACTGCAATGACAGAGTGT 4613
Db 1747 GCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGGAACAACTGCAATGACAGAGTGT 1806
QY 4614 GAGCCAGGAGCAGTACAGAGAGGTTGCTGGCTGAGAACACAGAGGCCACAGAGATGCGGAA 4673
Db 1807 GAGCCAGGAGCAGTACAGAGAGGTTGCTGGCTGAGAACACAGAGGCCACAGAGATGCGGAA 1866
QY 4674 GCTGTTTCAAGCTTGAGCCAGTCTCTGGGACCCGGAAGTGCAAGATGGACTCTTACCAAGCCCT 4733
Db 1867 GCTGTTTCAAGCTTGAGCCAGTCTCTGGGACCCGGAAGTGCAAGATGGACTCTTACCAAGCCCT 1926
QY 4734 GAAGGAGACCCATCCTCACTTATGGAATCTTGGGAGAGGGCAGCAAAAAGGACT 4793
Db 1927 GAAGGAGACCCATCCTCA - CTCACTTATGGAATCTTGGGAGAGGGCAGCAAAAAGGACT 1985
QY 4794 CTTGCCACTGAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTGAGTCTGCTGGCTTTGG 4853
Db 1986 CTTGCCACTGAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTTGAAGCTTGGCTTTGG 2045
QY 4854 CTGACCCCTTCTTTGGTCTCAGTTTCTTCTCTGCAAAAGTGGCCATCTGCTTTGGCT 4913
Db 2046 CTGACCCCTTCTTTGGTCTCAGTTTCTTCTCTGCAAAAGTGGCCATCTGCTTTGGCT 2105
QY 4914 TCCAGCATAAAGTAATGGAATCTTTGATGATGCTTTGCTGGGCAATTAATGTTCCATGCC 4973
Db 2106 TCCAGCATAAAGTAATGGAATCTTTGATGATGCTTTGCTGGGCAATTAATGTTCCATGCC 2165
QY 4974 AGGATGCGACAGGGGGCCCCAGTCCAGGTCCAGGTCCAGCATCTCAGGGAATGTCATC 5033
Db 2166 AGGATGCGACAGGGGGCCCCAGTCCAGGTCCAGGTCCAGCATCTCAGGGAATGTCATC 2225

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:06 ; Search time 1342.51 Seconds
(without alignments)
10946.442 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcctggagaggt.....aggaataagaagttacctac 5444

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5444	100.0	5523	22	AAF83651 Human CARD-7 polyp
2	5444	100.0	5523	25	AB556030 cDNA encoding huma
3	5439.2	99.9	5523	25	AB555497 cDNA encoding huma
4	4879.2	89.6	5100	24	AA479127 Pysin domain conta
5	4765.6	87.5	5122	22	AA012951 Human G-protein co
6	4148	76.2	4422	22	AA02760 Human NB-ARC and C
7	4078	74.9	4194	22	AA02761 Human NB-ARC and C
8	3952	72.6	4329	22	AA02762 Human NB-ARC and C

9	3400.4	62.5	4556	22	AA02764 Human NAC beta iso
10	3210.4	59.0	4466	22	AA02765 Human NAC gamma or
11	2572.6	47.3	2657	21	AAA78392 Human secreted pro
12	484.4	8.9	578	22	AA01744 Human reproductiv
13	484.4	8.9	578	23	ABL97037 Human testicular a
14	482.8	8.9	487	24	ABV97890 Human pancreatic c
15	414.4	7.6	416	24	ABV97891 Human pancreatic c
16	399.8	7.3	3186	24	AA44363 Human PYRIN-8 cDNA
17	398	7.3	3300	24	AA47129 Human PYRIN-11 cDN
18	369.8	6.8	3857	22	AA014323 Human PYRIN 1 (PYR
19	369.8	6.8	3857	25	ABX93556 Huma cDNA encoding
20	357.2	6.6	2763	24	ABQ75801 Human secreted pro
21	319.6	5.9	363	21	AAC08520 Human ORFX ORF1069
22	316.2	5.8	330	21	AAC75514 Human PYRIN-11 cDN
23	313.4	5.8	2691	24	AA44365 Human secreted pro
24	304.6	5.6	2847	22	AA01487 Secreted protein-e
25	304.6	5.6	2847	25	ABZ73494 Pysin domain conta
26	255.2	4.7	2199	24	AA47130 Human nucleotide b
27	254	4.7	2575	24	AA170683 NOVX related polyn
28	254	4.7	2767	25	ABT16018 Human secreted pro
29	254	4.7	3172	24	ABN99366 Human PYRIN-3 cDNA
30	252.4	4.6	3368	24	AA44366 Nucleotide sequenc
31	246.8	4.5	2614	25	ABV72512 Pysin domain conta
32	241.8	4.4	1557	24	AA47143 Human foetal liver
33	237	4.4	237	22	ABA69780 Human foetal liver
34	237	4.4	237	22	ABA36666 Probe #15132 for g
35	237	4.4	237	22	AAK17979 Human brain expres
36	237	4.4	237	22	AAK43859 Human bone marrow
37	237	4.4	237	22	AA149872 Probe #18558 used
38	237	4.4	237	22	ABS43511 Human liver single
39	237	4.4	237	24	ABS18088 Human genome-deriv
40	226.4	4.2	469	22	ABA58276 Human foetal liver
41	226.4	4.2	469	22	ABA27416 Probe #5882 for ge
42	226.4	4.2	469	22	AAK06371 Human brain expres
43	226.4	4.2	469	22	AAK32039 Human bone marrow
44	226.4	4.2	469	22	AA15917 Probe #5850 for ge
45	226.4	4.2	469	22	AA137897 Probe #6583 used t

ALIGNMENTS

RESULT 1
AAF83651
ID AAF83651 standard; cDNA; 5523 BP.

AC AAF83651;

DT 23-JUL-2001 (first entry)

DE Human CARD-7 polypeptide encoding cDNA.

XX CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;
KW autoimmune disorder; antiinflammatory; immunosuppressive; antiatherogenic;
KW antibacterial; antiviral; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
CDS 523..4812

FT /*tag= a

FT /product= "CARD-7"

XX WO200130813-A1.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-US29796.

XX PR 27-OCT-1999; 99US-0428252.

XX PA (MILL-) MILLENNIUM PHARM INC.

PI Bertin J;
 XX WPI: 2001-343340/36.
 DR P-PSDB; AAB62571.
 XX
 PT Identifying a modulator of interaction between caspase recruitment
 PT domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
 PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of
 PT the compound -
 XX
 PS Disclosure; Fig 1A-D; 80pp; English.
 XX
 CC The invention relates to identifying compounds, that modulate interaction
 CC between caspase recruitment domain (CARD)-7 and CARD-5. The method
 CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
 CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
 CC presence of the compound compared to the binding in the absence of the
 CC compound indicates that the compound is a modulator of CARD-7-CARD-5
 CC interaction). Modulators of CARD-7 and CARD-8 expression or activity can
 CC be used to treat or diagnose disorders such as cancers, bacterial or
 CC viral infections, autoimmune disorders (systemic lupus erythematosus,
 CC immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
 CC organ-specific autoimmunity including multiple sclerosis, Hashimoto's
 CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
 CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
 CC of cellular processes including cell growth and cell death. The present
 CC sequence represents a cDNA encoding the human CARD-7 polypeptide.
 XX
 SQ Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;
 Query Match 100.0%; Score 5444; DB 22; Length 5523;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCCAGGCGCTGAGAGGCTCTGAAGAAACCTGGAGAGCCAGCAGCCGGGGTCCACTCT 60
 DB 1 GCGCCAGGCGCTGAGAGGCTCTGAAGAAACCTGGAGAGCCAGCAGCCGGGGTCCACTCT 60
 QY 61 GGGTTCTGAAGCCCATTCCTGCTCTGGGCTCTCCACCCCTCTTCTCAGCCTT 120
 DB 61 GGGTTCTGAAGCCCATTCCTGCTCTGGGCTCTCCACCCCTCTTCTCAGCCTT 120
 QY 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGAGCCAGAGAGGAGAAATCTGAGGAA 180
 DB 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGAGCCAGAGAGGAGAAATCTGAGGAA 180
 QY 181 CAGAACAGTGAGGGTTGCCACACCCCATCTCCCGTCCACATCTCCCTCCACCTCAC 240
 DB 181 CAGAACAGTGAGGGTTGCCACACCCCATCTCCCGTCCACATCTCCCTCCACCTCAC 240
 QY 241 CCTCCCTGCTGGCCCTGAGACCCCATCCAGGAGCTCCCTATCAGTGACTTCTCCAGT 300
 DB 241 CCTCCCTGCTGGCCCTGAGACCCCATCCAGGAGCTCCCTATCAGTGACTTCTCCAGT 300
 QY 301 GTCTTGACGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCATCTCCCTCTAT 360
 DB 301 GTCTTGACGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCATCTCCCTCTAT 360
 QY 361 CGGCGTCTATCTGAGTCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
 DB 361 CGGCGTCTATCTGAGTCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
 QY 421 GACGTTAAGCCCAAGGAGAGAGCATGTTCTCTGCTGCTGCTGATACCTCACCC 480
 DB 421 GACGTTAAGCCCAAGGAGAGAGCATGTTCTCTGCTGCTGCTGATACCTCACCC 480
 QY 481 CTGGGAACATCCCCCAGACACCTCTTAATCTCGGAGACAGATGGCTGGCGAGCCTGG 540
 DB 481 CTGGGAACATCCCCCAGACACCTCTTAATCTCGGAGACAGATGGCTGGCGAGCCTGG 540
 QY 541 GCGCGCTGCGCTGTTTACTTGGAGTTCTGAGAGAGGAGGAGCTGAAGAGGTTCCAGCTT 600
 DB 541 GCGCGCTGCGCTGTTTACTTGGAGTTCTGAGAGAGGAGGAGCTGAAGAGGTTCCAGCTT 600

QY 601 CTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
 DB 601 CTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
 QY 661 AAGACAGTGGATGAGAGGTGGCTCTGTAATCTGTTGGCTCAGTATGGGAGAGCGGGCC 720
 DB 661 AAGACAGTGGATGAGAGGTGGCTCTGTAATCTGTTGGCTCAGTATGGGAGAGCGGGCC 720
 QY 721 TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGGTCACTGTCGCCCAAGCC 780
 DB 721 TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGGTCACTGTCGCCCAAGCC 780
 QY 781 CAGGAAGGGGAGGCGCACTCTCCCTCATTCCTTACAGCCCAAGTGAACCCACCTGGGG 840
 DB 781 CAGGAAGGGGAGGCGCACTCTCCCTCATTCCTTACAGCCCAAGTGAACCCACCTGGGG 840
 QY 841 TCTCCAGCAACCACTCTCCCTCATTCCTTACAGCCCAAGTGAACCCACCTGGGG 900
 DB 841 TCTCCAGCAACCACTCTCCCTCATTCCTTACAGCCCAAGTGAACCCACCTGGGG 900
 QY 901 GGGTGCACCCAGGCTCAGAGAGAGGGTTTGGACAGCTGCTTGACACATCTTGAGAGC 960
 DB 901 GGGTGCACCCAGGCTCAGAGAGAGGGTTTGGACAGCTGCTTGACACATCTTGAGAGC 960
 QY 961 CGTGGAGAGAAATCTCTGCTCTCTCTCTTACCAAGCTTCCCAAGCTCCCCAGACCAT 1020
 DB 961 CGTGGAGAGAAATCTCTGCTCTCTCTCTTACCAAGCTTCCCAAGCTCCCCAGACCAT 1020
 QY 1021 GAGTCTCAAGCCAGGAGTCAACCAAGCCCAATCCACAGCAGTCTGGGGAGCTGG 1080
 DB 1021 GAGTCTCAAGCCAGGAGTCAACCAAGCCCAATCCACAGCAGTCTGGGGAGCTGG 1080
 QY 1081 GGATCCCACTCAGCCAGCTTAGCACCACAGAGAGAGGGTCTCTGGGAGCCCAATGG 1140
 DB 1081 GGATCCCACTCAGCCAGCTTAGCACCACAGAGAGAGGGTCTCTGGGAGCCCAATGG 1140
 QY 1141 CCTCTGAGTGAAGCTCAGGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA 1200
 DB 1141 CCTCTGAGTGAAGCTCAGGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA 1200
 QY 1201 TCAGAAAGGAGCGCCCTCATGGGAGCGGTGTAGAAACGCCCCCAAGCGGCGACACC 1260
 DB 1201 TCAGAAAGGAGCGCCCTCATGGGAGCGGTGTAGAAACGCCCCCAAGCGGCGACACC 1260
 QY 1261 AGCTTAGAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACA 1320
 DB 1261 AGCTTAGAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACA 1320
 QY 1321 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTTACACAGCTGCTACTTCTACAAA 1380
 DB 1321 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTTACACAGCTGCTACTTCTACAAA 1380
 QY 1381 CCTCAACCCAGAGCCAGATCCCTTGTGTAAGAGAGCTGGCTGATTTATGTGAGAGAG 1440
 DB 1381 CCTCAACCCAGAGCCAGATCCCTTGTGTAAGAGAGCTGGCTGATTTATGTGAGAGAG 1440
 QY 1441 AATGAGAGACATTTAATTTGAGATCAGAGACTTTTGGCCCGAGGCTGGATACCCAGAA 1500
 DB 1441 AATGAGAGACATTTAATTTGAGATCAGAGACTTTTGGCCCGAGGCTGGATACCCAGAA 1500
 QY 1501 CCTCGCATAGTCATCTGACGGGGCTGCTGGAATTTGGAAAGTCAACACTGCCAGGCGAG 1560
 DB 1501 CCTCGCATAGTCATCTGACGGGGCTGCTGGAATTTGGAAAGTCAACACTGCCAGGCGAG 1560
 QY 1561 GTGAAGGAAGCTCTGGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCCAGCATGTCTTCTAC 1620
 DB 1561 GTGAAGGAAGCTCTGGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCCAGCATGTCTTCTAC 1620
 QY 1621 TTGAGCTGCGAGAGAGCTGGCCAGTCCAGTGGTGGTGGTCTCGCTGAGCTCATCGGAAAA 1680
 DB 1621 TTGAGCTGCGAGAGAGCTGGCCAGTCCAGTGGTGGTGGTCTCGCTGAGCTCATCGGAAAA 1680

QY 1681 GATGGAGAGCCTCCGGCTCCCATATTAGACAGATCCTGTCTAGGCGAGCGGCTGTCT 1740
 DB 1681 GATGGAGAGCCTCCGGCTCCCATATTAGACAGATCCTGTCTAGGCGAGCGGCTGTCT 1740
 QY 1741 TTCACTCTCGATGGGTAGATGAGCGCAGGATGGGTCTTGCAGAGCCGAGTTCTCAGCTC 1800
 DB 1741 TTCACTCTCGATGGGTAGATGAGCGCAGGATGGGTCTTGCAGAGCCGAGTTCTCAGCTC 1800
 QY 1801 TGTCTGCACTGGAGCCAGCCACAGCGCGGCGGATGCATGCTGCGGCGAGTTTGTGGGAAA 1860
 DB 1801 TGTCTGCACTGGAGCCAGCCACAGCGCGGCGGATGCATGCTGCGGCGAGTTTGTGGGAAA 1860
 QY 1861 ACTATACCTCCGAGGATCCTTCTGTATGACAGGCTCGGACACAGCTCTGAGAACCTC 1920
 DB 1861 ACTATACCTCCGAGGATCCTTCTGTATGACAGGCTCGGACACAGCTCTGAGAACCTC 1920
 QY 1921 ATTCTCTTTTGGAGCAGCAGCTGGGTAGAGGTCTTGGGTTCTCTGAGTCCAGCAGG 1980
 DB 1921 ATTCTCTTTTGGAGCAGCAGCTGGGTAGAGGTCTTGGGTTCTCTGAGTCCAGCAGG 1980
 QY 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG 2040
 DB 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG 2040
 QY 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTCTTGTGCTTGGGTTCTTGGGTTCTTGGGTTG 2100
 DB 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTCTTGTGCTTGGGTTCTTGGGTTCTTGGGTTG 2100
 QY 2101 TGCACCTGCTGATGAGCAGATGAAGCGGAGGAAAGCAATCAGCTGACTTCCAGACC 2160
 DB 2101 TGCACCTGCTGATGAGCAGATGAAGCGGAGGAAAGCAATCAGCTGACTTCCAGACC 2160
 QY 2161 ACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAAGCTCAGCCATTTGGAGCC 2220
 DB 2161 ACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAAGCTCAGCCATTTGGAGCC 2220
 QY 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCAATCTGGCAAAAAGACCTTTTC 2280
 DB 2221 CAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCAATCTGGCAAAAAGACCTTTTC 2280
 QY 2281 AGTCAGATGACCTCAGGAAGCATGGTTAGATGGGCTCATCTCCACCTTCTTGAG 2340
 DB 2281 AGTCAGATGACCTCAGGAAGCATGGGTTAGATGGGCTCATCTCCACCTTCTTGAG 2340
 QY 2341 ATGGGTATTTCTCAAGAGCACCCCATCTCTGAGCTACAGCTTCACTTCACTCTGTTC 2400
 DB 2341 ATGGGTATTTCTCAAGAGCACCCCATCTCTGAGCTACAGCTTCACTTCACTCTGTTC 2400
 QY 2401 CAAGAGTCTTTGACGCAATGTCTATGTCTTGGAGGATGAGAAGGGAGAGGTAAACAT 2460
 DB 2401 CAAGAGTCTTTGACGCAATGTCTATGTCTTGGAGGATGAGAAGGGAGAGGTAAACAT 2460
 QY 2461 TCTAATTGATCATAGATTTGGAAGACGCTAGAGCATATGGAATACATGGCCTGTTT 2520
 DB 2461 TCTAATTGATCATAGATTTGGAAGACGCTAGAGCATATGGAATACATGGCCTGTTT 2520
 QY 2521 GGGGATCAACACAGCTTTCTTATTTGGGCTGTTAAGTGTAGGGGAGAGAGATG 2580
 DB 2521 GGGGATCAACACAGCTTTCTTATTTGGGCTGTTAAGTGTAGGGGAGAGAGATG 2580
 QY 2581 GAGAACATCTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGAGTGGGTCCCGTCC 2640
 DB 2581 GAGAACATCTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGAGTGGGTCCCGTCC 2640
 QY 2641 CTGACGCTGTGTGAGCCACACTCTCTGAGGTCTCTGAGTCTCTGAGTCTGAGTCTGG 2700
 DB 2641 CTGACGCTGTGTGAGCCACACTCTCTGAGGTCTCTGAGTCTCTGAGTCTGAGTCTGG 2700
 QY 2701 AACAAAAGCTTCTGACACAGGTGATGGCCATTTTGAAGAAATGGGATGTGTAGAA 2760
 DB 2701 AACAAAAGCTTCTGACACAGGTGATGGCCATTTTGAAGAAATGGGATGTGTAGAA 2760
 QY 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTTAGCGCCACCGTGAAG 2820

DB 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTTGCATTAATTTAGCGCCACCGTGAAGAG 2820
 QY 2821 CTTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGGAGCCCGACCATGTTAGTCTG 2880
 DB 2821 CTTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGGAGCCCGACCATGTTAGTCTG 2880
 QY 2881 TTTAGGTTGGGTTCCAGTCAAGATGCTTATTTGGAGATTTCTTCTCCGCTCTCAAGGTC 2940
 DB 2881 TTTAGGTTGGGTTCCAGTCAAGATGCTTATTTGGAGATTTCTTCTCCGCTCTCAAGGTC 2940
 QY 2941 ACCAGAACCTTGAAGGAGCTGGAACCTAAGTGAAGAACTCGCTGAGGCACTCTGAGTGAAG 3000
 DB 2941 ACCAGAACCTTGAAGGAGCTGGAACCTAAGTGAAGAACTCGCTGAGGCACTCTGAGTGAAG 3000
 QY 3001 AGTCTTTTGAAGACCTTGAAGCGGCTCGCTGCTCTTGGAGACCTCGGTTGGCTGGC 3060
 DB 3001 AGTCTTTTGAAGACCTTGAAGCGGCTCGCTGCTCTTGGAGACCTCGGTTGGCTGGC 3060
 QY 3061 TGTGGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCAAACAGACC 3120
 DB 3061 TGTGGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCAAACAGACC 3120
 QY 3121 CTGACCGAGCTGGACCTTGAAGCTTCAATGTCTCAGGATCTGGAGCCAAACAGCTTTGC 3180
 DB 3121 CTGACCGAGCTGGACCTTGAAGCTTCAATGTCTCAGGATCTGGAGCCAAACAGCTTTGC 3180
 QY 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCAAGCTGCTGAGTGTGGCTTC 3240
 DB 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACTGCAAGCTGCTGAGTGTGGCTTC 3240
 QY 3241 ACCTCTGCTGCTGAGGACCTTGGCTTGTGCTTGTGCTGAGGCTGAGGCTGAGGAG 3300
 DB 3241 ACCTCTGCTGCTGAGGACCTTGGCTTGTGCTTGTGCTGAGGCTGAGGCTGAGGAG 3300
 QY 3301 CTGACCTTGGAGCAACAACTTGAATGAGTGTGGCTGCGACTGCTCTGTTGAGGGGCTC 3360
 DB 3301 CTGACCTTGGAGCAACAACTTGAATGAGTGTGGCTGCGACTGCTCTGTTGAGGGGCTC 3360
 QY 3361 AGGCATCTGCTGCAAACTCATACCTGCGGCTGGAACAGCAACAACTCTGAGTGTAG 3420
 DB 3361 AGGCATCTGCTGCAAACTCATACCTGCGGCTGGAACAGCAACAACTCTGAGTGTAG 3420
 QY 3421 ATGAGGAGGAACTGAGGGCTTGGAGCAGGAGAAACCTCAGCTGCTCATCTTACAGAGA 3480
 DB 3421 ATGAGGAGGAACTGAGGGCTTGGAGCAGGAGAAACCTCAGCTGCTCATCTTACAGAGA 3480
 QY 3481 CGGAAACCAAGTGTATGACCCCTACTGAGGCTTGGATACGGGAGAGATGAGTAAATAGC 3540
 DB 3481 CGGAAACCAAGTGTATGACCCCTACTGAGGCTTGGATACGGGAGAGATGAGTAAATAGC 3540
 QY 3541 ACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGGCTTCCATGCTGCTCAG 3600
 DB 3541 ACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGGCTTCCATGCTGCTCAG 3600
 QY 3601 GCTAATCTCAAACTCTTGAAGCTGAGCAAGATCTTCCCAATTTGCTGAGATTTCAGAGAA 3660
 DB 3601 GCTAATCTCAAACTCTTGAAGCTGAGCAAGATCTTCCCAATTTGCTGAGATTTCAGAGAA 3660
 QY 3661 AGCTCCCAGAGGATAGTACCGGTGAACTCTTGTGCTGCTTCTTCTGCTCTCAAGGG 3720
 DB 3661 AGCTCCCAGAGGATAGTACCGGTGAACTCTTGTGCTGCTTCTTCTGCTCTCAAGGG 3720
 QY 3721 GACCTGCAATGAGGCTTTGGGACTGAGTACTTCTGGGCTTCTGAGGCTTCTG 3780
 DB 3721 GACCTGCAATGAGGCTTTGGGACTGAGTACTTCTGGGCTTCTGAGGCTTCTG 3780
 QY 3781 GCTACTGAGGATAGTGAACAAAGAAAGAACTTGTACCGAGTTCACTTCTCTAGTCTGC 3840
 DB 3781 GCTACTGAGGATAGTGAACAAAGAAAGAACTTGTACCGAGTTCACTTCTCTAGTCTGC 3840
 QY 3841 TCCTACCGCTGCCCAACACCGGCTCTGCTTGTGATGAGAGAACGCTGACCGTTGAG 3900

Db 3841 TCCTACCGCTGGCCCAACACGGGTCTCTGCTTGTGATGAGAGAAAGCGGTGAG 3900
 QY 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 QY 3961 GTGGCAGGSCCTCTGTGTGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTTC 4020
 Db 3961 GTGGCAGGSCCTCTGTGTGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTTC 4020
 QY 4021 CCTCACTTGTGTGCTCTCAAGGGGCGCATGTGGACACATCCCTGTTCAAATGCGCCAC 4080
 Db 4021 CCTCACTTGTGTGCTCTCAAGGGGCGCATGTGGACACATCCCTGTTCAAATGCGCCAC 4080
 QY 4081 TTTAAAGAGAGGGGATGCTCTCGGAGAGCCAGCCAGGCTGAGCTGCATCACATAGTT 4140
 Db 4081 TTTAAAGAGAGGGGATGCTCTCGGAGAGCCAGCCAGGCTGAGCTGCATCACATAGTT 4140
 QY 4141 CTGGAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAAATGATCCATTAATGCCCTG 4200
 Db 4141 CTGGAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAAATGATCCATTAATGCCCTG 4200
 QY 4201 CGCTTCATTCCTGCTACCTCTGTGTGTGCTCTTACACCGGCTGCATCCTGAGGAATGC 4260
 Db 4201 CGCTTCATTCCTGCTACCTCTGTGTGTGCTCTTACACCGGCTGCATCCTGAGGAATGC 4260
 QY 4261 ACCTCCAGCTCTACCTGATCCCAAGTACTGCTCCATTCGGAAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTCCAGCTCTACCTGATCCCAAGTACTGCTCCATTCGGAAGAACTGGAGCTCTGC 4320
 QY 4321 TATCGAAGCCCTGGAGAAGACCAAGCTGTCTCGGAGTCTAGCTTGGCCACCTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAAGACCAAGCTGTCTCGGAGTCTAGCTTGGCCACCTTGGGATCA 4380
 QY 4381 GGGATCAGCTGCAAGTGAAGAACAAAGATGAGACTCTGTGTGGAGGCTTGGTG 4440
 Db 4381 GGGATCAGCTGCAAGTGAAGAACAAAGATGAGACTCTGTGTGGAGGCTTGGTG 4440
 QY 4441 AAACAGGAGATCTCATGCTCCCAAGTACTGATCCCTCCAGCCGATAGCCGTACCT 4500
 Db 4441 AAACAGGAGATCTCATGCTCCCAAGTACTGATCCCTCCAGCCGATAGCCGTACCT 4500
 QY 4501 TCACCTCTCGATGCCCCCGAGTGTCTGTCACCTTTGTGGACAGTATCGAGCAGCTGATA 4560
 Db 4501 TCACCTCTCGATGCCCCCGAGTGTCTGTCACCTTTGTGGACAGTATCGAGCAGCTGATA 4560
 QY 4561 GCGGAGTGACATCGGTGGAGTGTGTGGACAAACTGCATGAGCAGGTGTGAGCCAG 4620
 Db 4561 GCGGAGTGACATCGGTGGAGTGTGTGGACAAACTGCATGAGCAGGTGTGAGCCAG 4620
 QY 4621 GAGCAGTACGAGGGTGTGCTGAGACACGAGGCCAGCAGATCGGAGCTGTTTC 4680
 Db 4621 GAGCAGTACGAGGGTGTGCTGAGACACGAGGCCAGCAGATCGGAGCTGTTTC 4680
 QY 4681 AGCTTGAGCCAGTCTGTGGACCGGAAGTGCAGAAAGTGGACTTACCAAGCCCTGAAGAG 4740
 Db 4681 AGCTTGAGCCAGTCTGTGGACCGGAAGTGCAGAAAGTGGACTTACCAAGCCCTGAAGAG 4740
 QY 4741 ACCATCTCACTCATATTGAACTCTGGGAGAAAGGGCAGCAAAAGGATCTCTGCA 4800
 Db 4741 ACCATCTCACTCATATTGAACTCTGGGAGAAAGGGCAGCAAAAGGATCTCTGCA 4800
 QY 4801 CTCAGAGCTGAGTATCAACACAGCCCTTGACCCCTGAGTCTGCTGGCTTGGCTGACC 4860
 Db 4801 CTCAGAGCTGAGTATCAACACAGCCCTTGACCCCTGAGTCTGCTGGCTTGGCTGACC 4860
 QY 4861 TTCTTTGGGTCTCAGTCTCTCTCTGCAAAAGTGTGCAATCTGTTTGGCTTCCAGCA 4920
 Db 4861 TTCTTTGGGTCTCAGTCTCTCTCTGCAAAAGTGTGCAATCTGTTTGGCTTCCAGCA 4920
 QY 4921 CTAAGTAATGGAACCTTGTATGATGCTTGTGCGGCATTTATGTGCCATGCCAGGATG 4980
 Db 4921 CTAAGTAATGGAACCTTGTATGATGCTTGTGCGGCATTTATGTGCCATGCCAGGATG 4980

RESULT 2
 ABS56030

ID ABS56030 standard; cDNA; 5523 BP.

XX AC ABS56030;

XX DT 10-JAN-2003 (first entry)

XX DE cDNA encoding human caspase recruitment domain-7 (CARD-7).

XX KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudotICE; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 523..4812

XX FT /*tag= a

XX FT /product= "CARD-7"

XX PN US2002128198-A1.

XX PD 12-SEP-2002.

XX PF 27-NOV-2001; 2001US-0996617.

XX PR 28-JUN-1999; 99US-0340620.

XX PR 27-OCT-1999; 99US-0428252.

PR	15-AUG-2001; 2001US-0931071.	QY	301	GTCTTGAGGCCCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTCTAT	360
XX	(BERT/) BERTIN J.	DB	301	GTCTTGAGGCCCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTCTAT	360
XX	Bertin J;	QY	361	CGGCGTCTATCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
XX	WPI; 2003-028967/02.	DB	361	CGGCGTCTATCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
DR	P-PSDB; ABG71633.	QY	421	GACGGTAAGAGCAAGGCAAGGACAGCACTCTTCTGCTGCTGCTGATACCTCACCAC	480
XX	Identifying modulator of CARD-7 and CARD-5 interaction, by contacting	DB	421	GACGGTAAGAGCAAGGCAAGGACAGCACTCTTCTGCTGCTGCTGATACCTCACCAC	480
PT	CARD-7 and CARD-5 in presence of test compound, measuring their	QY	481	CTGGGAACATCCCCCAGACACCCCTCTTAACCTCGGAGACAGATGGCTGGGAGGCTGG	540
PT	binding, and identifying modulator, when binding of CARD-7 to CARD-5 is	DB	481	CTGGGAACATCCCCCAGACACCCCTCTTAACCTCGGAGACAGATGGCTGGGAGGCTGG	540
XX	altered	QY	541	GGCCGCTGGGCTGTACTTGGAGTTCTTGAGAGGAGGAGCTGAAGAGTTCCAGCTT	600
PS	Disclosure; Fig 1; 43pp; English.	DB	541	GGCCGCTGGGCTGTACTTGGAGTTCTTGAGAGGAGGAGCTGAAGAGTTCCAGCTT	600
XX	The present invention relates to methods of identifying compounds	QY	601	CTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG	660
CC	that regulate caspase activity using caspase recruitment domain-7	DB	601	CTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG	660
CC	(CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a	QY	661	AAGACAGTGGGATGGAGGTGGCTCTGACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
CC	method for identifying a compound that modulates the interaction	DB	661	AAGACAGTGGGATGGAGGTGGCTCTGACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
CC	between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta	QY	721	TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
CC	converting enzyme (pseudolICE) or ICEBERG is disclosed. CARD-7 and	DB	721	TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
CC	CARD-8 molecules are useful as modulating agents in regulating a	QY	781	CAGGAAGGCGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCCACTGGGG	840
CC	variety of cellular processes including cell growth, cell death, and	DB	781	CAGGAAGGCGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCCACTGGGG	840
CC	inflammation. The methods of the invention are useful for identifying	QY	841	TCTCCAGCAACCCCACTCCACCGCAGTGTAAATGCTCGATCATGAATGCGGGG	900
CC	compounds that have the ability to increase/decrease apoptosis, or	DB	841	TCTCCAGCAACCCCACTCCACCGCAGTGTAAATGCTCGATCATGAATGCGGGG	900
CC	comprise the ability to induce caspase activation. The methods are	QY	901	GGGTGACCCAGGCTCAGAGAGAGGTTTGGACAGCTGCCTGACACATCTGGACCC	960
CC	useful for treating a disorder associated with inappropriate apoptosis	DB	901	GGGTGACCCAGGCTCAGAGAGAGGTTTGGACAGCTGCCTGACACATCTGGACCC	960
CC	or inappropriate inflammation. The methods are useful for treating	QY	961	CGCTGGAGAGAAATCTCTGCTCTCACTCTCACTCAAGCTCTTCCAAAGCTTCCCAAGCTT	1020
CC	disorders associated with an undesirably low rate of apoptosis such	DB	961	CGCTGGAGAGAAATCTCTGCTCTCACTCTCACTCAAGCTCTTCCAAAGCTTCCCAAGCTT	1020
CC	as cancer (preferably follicular lymphoma, chronic myelogenous	QY	1021	GAGTCTCCAAGCCAGGAGTCACTCCCAAGCCCAATCCACAGAGTGTCTGGGAGCTGG	1080
CC	leukemia, melanoma, colon cancer, lung carcinoma, etc), viral	DB	1021	GAGTCTCCAAGCCAGGAGTCACTCCCAAGCCCAATCCACAGAGTGTCTGGGAGCTGG	1080
CC	infections, autoimmune diseases caused by low levels of apoptosis	QY	1081	GGATCCCACTCAGCCAGCTTAGCACCCAGAGCAGGAGCTCTCTGGGAGCTGG	1140
CC	(e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,	DB	1081	GGATCCCACTCAGCCAGCTTAGCACCCAGAGCAGGAGCTCTCTGGGAGCTGG	1140
CC	and arthritis). The methods are also useful for treating disorders with	QY	1141	CTCTGGATGAAACGTCAGGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
CC	undesirably high rates of apoptosis such as human immunodeficiency	DB	1141	CTCTGGATGAAACGTCAGGAAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
CC	virus (HIV) infection, Alzheimer's disease, Parkinson's disease,	QY	1201	TCAGAGAAAGGAGGCGCCCTCAGCCAGCTTAGCACCCAGAGCAGGAGCTCTCTGGGAGCTGG	1260
CC	amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal	DB	1201	TCAGAGAAAGGAGGCGCCCTCAGCCAGCTTAGCACCCAGAGCAGGAGCTCTCTGGGAGCTGG	1260
CC	muscular atrophy, various forms of cerebellar degeneration, anemia	QY	1261	AGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACCA	1320
CC	associated with chronic disease, aplastic anemia, chronic neutropenia,	DB	1261	AGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACCA	1320
CC	myelodysplastic syndromes, myocardial infarction, stroke, and	QY	1321	TGGCCCTGAAAGAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTTCTTACAAA	1380
CC	various inflammatory disorders (e.g. Crohn's disease, reactive	DB	1321	TGGCCCTGAAAGAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTTCTTACAAA	1380
CC	arthritis, insulin dependent diabetes mellitus, food allergies,	QY	1381	CCTCACCCCAAGCAAGATCCCTGGTCAAGAGAGAGTGGCCCTGATTTATGTGGAGGAG	1440
CC	psoriasis, graft rejection, allergic rhinitis, multiple sclerosis,	DB			
CC	conjunctivitis, glomerular nephritis, etc). The present sequence				
CC	encodes human CARD-7.				
XX					
SQ	Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;				
	Query Match 100.0%; Score 5444; DB 25; Length 5523;				
	Best Local Similarity 100.0%; Pred. No. 0;				
	Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 GCCCCAGGCGCTGAGAGGTCTGAAGAAACCTGGAGCAGAGCCCGGGCTCCACTCT	60			
DB	1 GCCCCAGGCGCTGAGAGGTCTGAAGAAACCTGGAGCAGAGCCCGGGCTCCACTCT	60			
QY	61 GGGTCTGAAGCCCATTCCTGCTGCGGCTCTCCCAACCCCACTCTTCTCAGCCTT	120			
DB	61 GGGTCTGAAGGCCCATTCCTGCTGCGGCTCTCCCAACCCCACTCTTCTCAGCCTT	120			
QY	121 GCAGCTCAAGGGTTGATCTCAGGAGTCCAGAGCCAGGAGGGAAGAATCTCAGGAACA	180			
DB	121 GCAGCTCAAGGGTTGATCTCAGGAGTCCAGAGCCAGGAGGGAAGAATCTCAGGNACA	180			
QY	181 CAGAACAGTGAAGGTTGCCACACCCCATCTCCCGTCAACACATCTCCCTCCACCTCAC	240			
DB	181 CAGAACAGTGAAGGTTGCCACACCCCATCTCCCGTCAACACATCTCCCTCCACCTCAC	240			
QY	241 CCTCCCTGCTGGCCCTGGACCCCATCCAGGAGCTTCCCTATCAGTGACTTCTCCAGT	300			
DB	241 CCTCCCTGCTGGCCCTGGACCCCATCCAGGAGCTTCCCTATCAGTGACTTCTCCAGT	300			

[illegible]

Db	2461	TCTAAATTGGCATCATAGATTGTTGGAAAAAGACGCTAGAGAGCATATGGAATACATGAGCCTGTGTTT	2520
Qy	2521	GGGCGATCAACCAACACACGTTTCTTAATTGGGCGCTGTTAAGTGCATGAGGGGGAGAGAGAGATG	2580
Db	2521	GGGCGATCAACCAACACGTTTCTTAATTGGGCGCTGTTAAGTGCATGAGGGGGAGAGAGAGATG	2580
Qy	2581	GAGAAACATCTTTTCACTGCGCGCTGTCTCAGGGAGGAAACCTGATGCAAGTGGGTCTCCGTCC	2640
Db	2581	GAGAAACATCTTTTCACTGCGCGCTGTCTCAGGGAGGAAACCTGATGCAAGTGGGTCTCCGTCC	2640
Qy	2641	CTGCAGCTGCTGTGCAGCCACACACTCTCTGGAGTCCCTCCACTGCTGTGTACGAGACTCGG	2700
Db	2641	CTGCAGCTGCTGTGCAGCCACACACTCTCTGGAGTCCCTCCACTGCTGTGTACGAGACTCGG	2700
Qy	2701	AACAAAAAGTTTCTTGACACAAGTGTAGGGCCATTTTCGAAGAAATGGGCATGTGTGTAGAA	2760
Db	2701	AACAAAAAGTTTCTTGACACAAGTGTAGGGCCATTTTCGAAGAAATGGGCATGTGTGTAGAA	2760
Qy	2761	ACAGACATGGAGCTCTTTPAGTGTGCATTTTCTGCATTTAAATTCAGGCCCAACGTGAAGAAG	2820
Db	2761	ACAGACATGGAGCTCTTTPAGTGTGCATTTTCTGCATTTAAATTCAGGCCCAACGTGAAGAAG	2820
Qy	2821	CTTCAGCTGATTGAGGCGAGCGACACAGATCAACATGGAGCCCAACCATGGTGTAGTCCCTG	2880
Db	2821	CTTCAGCTGATTGAGGCGAGCGACACAGATCAACATGGAGCCCAACCATGGTGTAGTCCCTG	2880
Qy	2881	TTCAAGTGGGTCCACAGTCACAGATGCCATTGTGGCAGATTCTCTTCTCCGTCCTCAAGGTC	2940
Db	2881	TTCAAGTGGGTCCACAGTCACAGATGCCATTGTGGCAGATTCTCTTCTCCGTCCTCAAGGTC	2940
Qy	2941	ACCAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCAGTGAAG	3000
Db	2941	ACCAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCAGTGAAG	3000
Qy	3001	AGTCTTTTGAAGACCTGAGACGCCCTCGCTGCTCCTGTGAGACACCTGCGGTTGGCTGCG	3060
Db	3001	AGTCTTTTGAAGACCTGAGACGCCCTCGCTGCTCCTGTGAGACACCTGCGGTTGGCTGCG	3060
Qy	3061	TGTGGCCTCACAGCTGAGGACTCGAAGGACCTTGGCTTTGGGCTGAGAGCAACACAGACC	3120
Db	3061	TGTGGCCTCACAGCTGAGGACTCGAAGGACCTTGGCTTTGGGCTGAGAGCAACACAGACC	3120
Qy	3121	CTGACCGAGCTGGAACCTTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTGC	3180
Db	3121	CTGACCGAGCTGGAACCTTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTGC	3180
Qy	3181	CAGAGACTGAGACAGCCGAGCTGGAAGCTACAGGACTGCACTGTGTGAGCTGTGGCCTC	3240
Db	3181	CAGAGACTGAGACAGCCGAGCTGGAAGCTACAGGACTGCACTGTGTGAGCTGTGGCCTC	3240
Qy	3241	ACGTCCTACCTGTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCGCCCTGAAGGAG	3300
Db	3241	ACGTCCTACCTGTGCCAGGACCTGGCTCTGTGCTTAGTGCCAGCCGCCCTGAAGGAG	3300
Qy	3301	CTAGACCTGACAGAACAACTCGGATGACGTTTGGCTGCACTGCTGTGAGGGGCTC	3360
Db	3301	CTAGACCTGACAGAACAACTCGGATGACGTTTGGCTGCACTGCTGTGAGGGGCTC	3360
Qy	3361	AGGCATCTCGCTGCAAACTCATACGCTGGGCTGGAACAGACAACACTGTAGTGTATGAG	3420
Db	3361	AGGCATCTCGCTGCAAACTCATACGCTGGGCTGGAACAGACAACACTGTAGTGTATGAG	3420
Qy	3421	ATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAAACCTCAGCTGCTCATCTTTCAGCAGA	3480
Db	3421	ATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAAACCTCAGCTGCTCATCTTTCAGCAGA	3480
Qy	3481	CGGAAACCAAGTGTGATGACCCCTACTGTAGGGGCTCGATACGGGAGAGATGAGTAATAGC	3540
Db	3481	CGGAAACCAAGTGTGATGACCCCTACTGTAGGGGCTCGATACGGGAGAGATGAGTAATAGC	3540
Qy	3541	ACATCTCATCTAAGCGGCGAGACTCGGATCAGAGAGGGCGGCTCCCATGTTGCTCAG	3600
Db	3541	ACATCTCATCTAAGCGGCGAGACTCGGATCAGAGAGGGCGGCTCCCATGTTGCTCAG	3600

QY 3601 GCTAATCTCAAACTCTCGACGCTGACAAAGATCTTCCCAATTCGTGAGATTCGAGGAA 3660
 Db 3601 GCTAATCTCAAACTCTCGACGCTGACAAAGATCTTCCCAATTCGTGAGATTCGAGGAA 3660
 QY 3661 AGCTCCCGAGAGGTAGTACGGTGGAACTCTTGTGGTGGTCTCTCTGCTCTCAAGG 3720
 Db 3661 AGCTCCCGAGAGGTAGTACGGTGGAACTCTTGTGGTGGTCTCTCTGCTCTCAAGG 3720
 QY 3721 GACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGCTCTG 3780
 Db 3721 GACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGCTCTG 3780
 QY 3781 GCTACTGAGGTAGTTGCAAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 Db 3781 GCTACTGAGGTAGTTGCAAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 QY 3841 TCCCTACCGTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAGAACGGTGTGAG 3900
 Db 3841 TCCCTACCGTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAGAACGGTGTGAG 3900
 QY 3901 ATTGAATCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 QY 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGAACTGTGCACTTC 4020
 Db 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGAACTGTGCACTTC 4020
 QY 4021 CCTCACTTTGTGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTTCAAAATGGCCAC 4080
 Db 4021 CCTCACTTTGTGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTTCAAAATGGCCAC 4080
 QY 4081 TTTTAAAGAGAGGGAGTGTCTCTGGAGAGCCAGCCAGGCTGAGCTGCATCACATAGTT 4140
 Db 4081 TTTTAAAGAGAGGGAGTGTCTCTGGAGAGCCAGCCAGGCTGAGCTGCATCACATAGTT 4140
 QY 4141 CTGGAACACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAATGATCATATGCCCCTG 4200
 Db 4141 CTGGAACACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAATGATCATATGCCCCTG 4200
 QY 4201 CGCTTCATCTCCGCTACCTCTGTGTGTGTCTTTCACCGGCTGCATCTGAGGAAGTC 4260
 Db 4201 CGCTTCATCTCCGCTACCTCTGTGTGTGTCTTTCACCGGCTGCATCTGAGGAAGTC 4260
 QY 4261 ACCTTCACCTCTACCTGATCCCAAGTCTGCTGCTCCATTCGGAAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTTCACCTCTACCTGATCCCAAGTCTGCTGCTCCATTCGGAAGAACTGGAGCTCTGC 4320
 QY 4321 TATCGAAGCCCTGGAGAAGACCAAGCTGTTCTCGAGTTCACGTTGGCCACTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAAGACCAAGCTGTTCTCGAGTTCACGTTGGCCACTTGGGATCA 4380
 QY 4381 GGGATCAGGCTGCAAGTGAAGACCAAGAGATGAGACTCTGGTGTGGAGGCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGACCAAGAGATGAGACTCTGGTGTGGAGGCTTGGTG 4440
 QY 4441 AAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCGATAGCCGTACCT 4500
 Db 4441 AAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCGATAGCCGTACCT 4500
 QY 4501 TCACCTCTGATGCCCGCAGTGTGCTGCACTTTGTGGACAGTATCGAGACAGCTGATA 4560
 Db 4501 TCACCTCTGATGCCCGCAGTGTGCTGCACTTTGTGGACAGTATCGAGACAGCTGATA 4560
 QY 4561 GCCCGAGTGCATCGGTGGAGGTGTCTTGGACAACTGCATGCAGACAGTGTGAGCCAG 4620
 Db 4561 GCCCGAGTGCATCGGTGGAGGTGTCTTGGACAACTGCATGCAGACAGTGTGAGCCAG 4620
 QY 4621 GACAGTACAGAGGTGTGCTGTGAGAACACAGAGCCAGCCAGATCGGGAAGCTGTTC 4680
 Db 4621 GACAGTACAGAGGTGTGCTGTGAGAACACAGAGCCAGCCAGATCGGGAAGCTGTTC 4680

QY 4681 AGCTTGAGCCAGTCTCTGGACCGGAGTGCAGAGATGAGACTCTACCAAGCCCTGAGGAG 4740
 Db 4681 AGCTTGAGCCAGTCTCTGGACCGGAGTGCAGAGATGAGACTCTACCAAGCCCTGAGGAG 4740
 QY 4741 ACCCATCTCACTCATTTATGAACTCTGGGAGAGGCGAGCAAAAGAGGACTCTCTGCCA 4800
 Db 4741 ACCCATCTCACTCATTTATGAACTCTGGGAGAGGCGAGCAAAAGAGGACTCTCTGCCA 4800
 QY 4801 CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTTGAGTCTCTGGCTTGGCTGACCC 4860
 Db 4801 CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTTGAGTCTCTGGCTTGGCTGACCC 4860
 QY 4861 TTTCTTTGGTCTCAGTCTTCTCTGCAACAAAGTGTGCCATCTGCTTGGCTTCCAGCA 4920
 Db 4861 TTTCTTTGGTCTCAGTCTTCTCTGCAACAAAGTGTGCCATCTGCTTGGCTTCCAGCA 4920
 QY 4921 CTAAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4980
 Db 4921 CTAAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4980
 QY 4981 CCACAGGGGGCCCCAGTCCAGTGGCTTAAACAGCATCTCAGGGAATGTCTCTGGAGCT 5040
 Db 4981 CCACAGGGGGCCCCAGTCCAGTGGCTTAAACAGCATCTCAGGGAATGTCTCTGGAGCT 5040
 QY 5041 GGCAGAGACCCCTGAGACCTCTATAGAGCTCTATCTGGTGGCCACAGAGCCCAAGCTAGA 5100
 Db 5041 GGCAGAGACCCCTGAGACCTCTATAGAGCTCTATCTGGTGGCCACAGAGCCCAAGCTAGA 5100
 QY 5101 GGCCTCCGGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCAATTTGCCTC 5160
 Db 5101 GGCCTCCGGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCAATTTGCCTC 5160
 QY 5161 TGCTGTGTGTCAGAGGCTGAGCCGCCAAATTTGGGGTTTTCAGCGTGGAGGCGCAGTGGATT 5220
 Db 5161 TGCTGTGTGTCAGAGGCTGAGCCGCCAAATTTGGGGTTTTCAGCGTGGAGGCGCAGTGGATT 5220
 QY 5221 TTGGCTTTGTACAGGAAGTCTACAGAGCAAGCAACAGAGTAAGTGGAGGAAGATT 5280
 Db 5221 TTGGCTTTGTACAGGAAGTCTACAGAGCAAGCAACAGAGTAAGTGGAGGAAGATT 5280
 QY 5281 ATTACAGAAATAAGAGGATATCACAGCTCTTTTAAATTTTAAATTTTAAATTTTAAATTT 5340
 Db 5281 ATTACAGAAATAAGAGGATATCACAGCTCTTTTAAATTTTAAATTTTAAATTTTAAATTT 5340
 QY 5341 TTTTACAGAAACCCCTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 5400
 Db 5341 TTTTACAGAAACCCCTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 5400
 QY 5401 ACAAAGAAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 5444
 Db 5401 ACAAAGAAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 5444

RESULT 3
 ABS5497
 ID ABS5497 standard; cdna; 5523 BP.
 XX
 AC ABS5497;
 XX
 DX 09-JAN-2003 (first entry)
 XX
 DE cdna encoding human caspase recruitment domain-7 (CARD-7).
 XX
 KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudoleukin-1 beta converting enzyme; IL-1beta;
 KW pseudolice; ICSEBER; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; ALS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;

KW	graft rejection; allergic rhinitis; food allergy; conjunctivitis;	QY	61	GGTTTCTGAAAGCCCATTCCTGCTCTCGGCTCTCCACCCACCTCTTCTCAGCCTT	120
KW	glomerular nephritis; cytostatic; virucide; immunosuppressive;	Db	61	GGTTTCTGAAAGCCCATTCCTGCTCTCGGCTCTCCACCCACCTCTTCTCAGCCTT	120
XX	dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.				
OS	Homo sapiens.				
XX					
XX	Location/Qualifiers				
FT	523..4812				
FT	/tag= a				
FT	/product= "CARD-7"				
XX					
XX	US2002128219-A1.				
PD	12-SEP-2002.				
XX					
XX	15-AUG-2001; 2001US-0931071.				
XX					
PR	27-OCT-1999; 99US-0428252.				
XX					
PA	(BERT/) BERTIN J.				
PA	(ALNE/) ALNEMRI E. S.				
XX					
XX	Bertin J, Alnemri ES;				
PI	WPI; 2003-028968/02.				
XX	P-PSDB; ABG71631.				
DR					
XX					
PT	Assays for identifying compound that modulates the interaction of				
PT	caspase recruitment domain-8 with a CARD-8 ligand or a compound that				
PT	modulates activity of CARD-8 -				
XX					
PS	Disclosure; Fig 1; 49pp; English.				
XX					
CC	The present invention relates to methods of identifying compounds				
CC	that regulate caspase activity using caspase recruitment domain-7				
CC	(CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a				
CC	method for identifying a compound that modulates the interaction				
CC	between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta				
CC	converting enzyme (pseudolice) or ICEBERG is disclosed. CARD-7 and				
CC	CARD-8 molecules are useful as modulating agents in regulating a				
CC	variety of cellular processes including cell growth, cell death, and				
CC	inflammation. The methods of the invention are useful for identifying				
CC	compounds that have the ability to increase/decrease apoptosis, or				
CC	comprise the ability to induce caspase activation. The methods are				
CC	useful for treating a disorder associated with inappropriate apoptosis				
CC	or inappropriate inflammation. The methods are useful for treating				
CC	disorders associated with an undesirably low rate of apoptosis such				
CC	as cancer (preferably follicular lymphoma, chronic myelogenous				
CC	leukemia, melanoma, colon cancer, lung carcinoma, etc), viral				
CC	infections, autoimmune diseases caused by low levels of apoptosis				
CC	(e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,				
CC	and arthritis). The methods are also useful for treating disorders with				
CC	undesirably high rates of apoptosis such as human immunodeficiency				
CC	virus (HIV) infection, Alzheimer's disease, Parkinson's disease,				
CC	amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal				
CC	muscular atrophy, various forms of cerebellar degeneration, anaemia				
CC	associated with chronic disease, aplastic anaemia, chronic neutropenia,				
CC	myelodysplastic syndromes, myocardial infarction, stroke, and				
CC	various inflammatory disorders (e.g. Crohn's disease, reactive				
CC	arthritis, insulin dependent diabetes mellitus, multiple sclerosis,				
CC	psoriasis, graft rejection, allergic rhinitis, food allergies,				
CC	conjunctivitis, glomerular nephritis, etc). The present sequence				
CC	encodes human CARD-7.				
XX					
SQ	Sequence 5523 BP; 1350 A; 1529 C; 1481 G; 1163 T; 0 other;				
	Query Match 99.9%; Score 5439.2; DB 25; Length 5523;				
	Best Local Similarity 99.9%; Pred. No. 0;				
	Matches 5441; Conservative 0; Mismatches 3; Indels 0; Gaps 0;				
QY	1 GCCCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGACGCCGGGGCTCACTCT				60
Db	1 GCCCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGACGCCGGGGCTCACTCT				60

QY	61	GGTTTCTGAAAGCCCATTCCTGCTCTCGGCTCTCCACCCACCTCTTCTCAGCCTT	120
Db	61	GGTTTCTGAAAGCCCATTCCTGCTCTCGGCTCTCCACCCACCTCTTCTCAGCCTT	120
QY	121	GCAGCTCAAGGGTTGATCTCAGGAGTCCAGAGCCAGAGAGGGAAGATCTGAGGAACA	180
Db	121	GCAGCTCAAGGGTTGATCTCAGGAGTCCAGAGCCAGAGAGGGAAGATCTGAGGAACA	180
QY	181	CAGAACAGTGAGCGTTGCCACACCCCATCTCCCGTCCACCATCTCCCTCACCCTC	240
Db	181	CAGAACAGTGAGCGTTGCCACACCCCATCTCCCGTCCACCATCTCCCTCACCCTC	240
QY	241	CCTCCCTGCTGGCCCTGGAGCCCATCCAGAGACCTCCCTATCAGCTGACTTCTTCCAGT	300
Db	241	CCTCCCTGCTGGCCCTGGAGCCCATCCAGAGACCTCCCTATCAGCTGACTTCTTCCAGT	300
QY	301	GTCTTGCAGGCGCTCTGGGCTCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTAT	360
Db	301	GTCTTGCAGGCGCTCTGGGCTCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTAT	360
QY	361	CGGCGTCTATCTGTAGTGCCTGGGATTTATAAATCTGGGTTCCGAATCTGGAATAAGA	420
Db	361	CGGCGTCTATCTGTAGTGCCTGGGATTTATAAATCTGGGTTCCGAATCTGGAATAAGA	420
QY	421	GACGGTAAGAGCCAAAGGCAAGGACAGCACTGTTCTCTGCTGCTGCTGATACCTCACCAC	480
Db	421	GACGGTAAGAGCCAAAGGCAAGGACAGCACTGTTCTCTGCTGCTGCTGATACCTCACCAC	480
QY	481	CTGGGAAACATCCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGGGAGCTGG	540
Db	481	CTGGGAAACATCCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGGGAGCTGG	540
QY	541	GGCGGCTGGCTGTACTTGGAGTTCTGAGAGGAGGAGCTGAAGAGTTCAGCTT	600
Db	541	GGCGGCTGGCTGTACTTGGAGTTCTGAGAGGAGGAGCTGAAGAGTTCAGCTT	600
QY	601	CTGCTCGGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG	660
Db	601	CTGCTCGGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCTGCTCAGCCAGAG	660
QY	661	AAGACAGTGGATGAGAGTGGCTCTGTAACCTGCTGCTGCTGATGAGGAGAGCGGGCC	720
Db	661	AAGACAGTGGATGAGAGTGGCTCTGTAACCTGCTGCTGCTGATGAGGAGAGCGGGCC	720
QY	721	TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGTGCGCCCAAGCC	780
Db	721	TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGTGCGCCCAAGCC	780
QY	781	CAGGAAGGGGAGGAGGAGCTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCACCTGGGG	840
Db	781	CAGGAAGGGGAGGAGGAGCTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCACCTGGGG	840
QY	841	TCTCCAGCCAAACCCATCTCCAGCGAGTGTATATGCTGCTGATCCATGATGCGGGG	900
Db	841	TCTCCAGCCAAACCCATCTCCAGCGAGTGTATATGCTGCTGATCCATGATGCGGGG	900
QY	901	GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGGAGCG	960
Db	901	GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGGAGCG	960
QY	961	CGCTGGAGAGAAATCTCTGCTCTCACTCTCTTACCAAGCTCTTCAAGCTTCCCAAGC	1020
Db	961	CGCTGGAGAGAAATCTCTGCTCTCACTCTCTTACCAAGCTCTTCAAGCTTCCCAAGC	1020
QY	1021	GAGTCTCAAGCCAGGAGTCAACCAACCGCCCTCCATCCACAGCAGTCTGGGAGCTGG	1080
Db	1021	GAGTCTCAAGCCAGGAGTCAACCAACCGCCCTCCATCCACAGCAGTCTGGGAGCTGG	1080
QY	1081	GGATCCCCACTCAGCCCGCTAGCAGCCAGAGAGGAGGCTCTCTGGGAGCCCAATGG	1140
Db	1081	GGATCCCCACTCAGCCCGCTAGCAGCCAGAGAGGAGGCTCTCTGGGAGCCCAATGG	1140

Db 3301 CTAGACCTGCAGCAGAACCAACCTGGATGACGTTGGCGTGCAGCTGCTGTGAGGGGCTC 3360
 Qy 3361 AGGCATCTCTGCTGCAAACTCATACGCCCTGGGCTGCAGACAGAACAACTCTGAGTGTATGAG 3420
 Db 3361 AGGCATCTCTGCTGCAAACTCATACGCCCTGGGCTGCAGACAGAACAACTCTGAGTGTATGAG 3420
 Qy 3421 ATGAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480
 Db 3421 ATGAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480
 Qy 3481 CGGAAACCAAGTGTATGATACCCCTCATGAGGGCTTGATACGGAGAGATGATGATATAGC 3540
 Db 3481 CGGAAACCAAGTGTATGATACCCCTCATGAGGGCTTGATACGGAGAGATGATGATATAGC 3540
 Qy 3541 ACATCCTCACTCAAGCGGAGAGACTCGATCAGAGAGGGCGCTTCCCATGTTGCTCAG 3600
 Db 3541 ACATCCTCACTCAAGCGGAGAGACTCGATCAGAGAGGGCGCTTCCCATGTTGCTCAG 3600
 Qy 3601 GCTAATCTCAAACTCTCGACGCTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGAA 3660
 Db 3601 GCTAATCTCAAACTCTCGACGCTGAGCAAGATCTTCCCAATTGCTGAGATTGCAGAGAA 3660
 Qy 3661 AGCTCCCAAGAGTAGTAACGGTGGAACTCTTGTGCGTGCCTTCTCCTGCCCTCTCAAGGG 3720
 Db 3661 AGCTCCCAAGAGTAGTAACGGTGGAACTCTTGTGCGTGCCTTCTCCTGCCCTCTCAAGGG 3720
 Qy 3721 GACCTGATCAAGAGCTTGTGGGACTGACGATGACTTCTGGGGCCCAAGGGGCTGTG 3780
 Db 3721 GACCTGATCAAGAGCTTGTGGGACTGACGATGACTTCTGGGGCCCAAGGGGCTGTG 3780
 Qy 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTCTGAGCTGGC 3840
 Db 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTTCACTTCCCTCTGAGCTGGC 3840
 Qy 3841 TCTACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Db 3841 TCTACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Qy 3901 ATTGAATCTGTGTGGAGCAGTTCTTGGGTGAGATCAACCCACAGCAGCAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGAGCAGTTCTTGGGTGAGATCAACCCACAGCAGCAGCTGGATG 3960
 Qy 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGAGCTGTGGAAGCTGTGACACCTC 4020
 Db 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGAGCTGTGGAAGCTGTGACACCTC 4020
 Qy 4021 CCTCACTTTGTGGCTCTCAAGGGGCCATGTGGAACATCTCCTGTTCAAAATGGCCAC 4080
 Db 4021 CCTCACTTTGTGGCTCTCAAGGGGCCATGTGGAACATCTCCTGTTCAAAATGGCCAC 4080
 Qy 4081 TTTTAAGAGGGGATGCTCTGGAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 Db 4081 TTTTAAGAGGGGATGCTCTGGAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 Qy 4141 CTGGAAAAACCCAGCTTCTCCCTTGGAGTCTCCTCGAAATGATCATTAATGCCCTG 4200
 Db 4141 CTGGAAAAACCCAGCTTCTCCCTTGGAGTCTCCTCGAAATGATCATTAATGCCCTG 4200
 Qy 4201 CGCTTCATTTCCGCTCACTCTGTGTGTTGCTTTTACACCGGCTGCATCCTGAGGAAGTC 4260
 Db 4201 CGCTTCATTTCCGCTCACTCTGTGTGTTGCTTTTACACCGGCTGCATCCTGAGGAAGTC 4260
 Qy 4261 ACCTTCCACTCTACCTGATCCCAAGTACTGCTCCATTCGGAGAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTTCCACTCTACCTGATCCCAAGTACTGCTCCATTCGGAGAGAACTGGAGCTCTGC 4320
 Qy 4321 TATCGAAGCCCTGGAGAGAACAGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAGAACAGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCA 4380
 Qy 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTCTGGTGTGGAGGCGCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTCTGGTGTGGAGGCGCTTGGTG 4440

RESULT 4
 AAL47127
 ID AAL47127 standard; DNA; 5100 BP.

Qy 4441 AAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCAGCCCGCATAGCCGTACTT 4500
 Db 4441 AAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCAGCCCGCATAGCCGTACTT 4500
 Qy 4501 TCACCTCTGGATGCCCCCGAGTTGCTGCACTTTGTGGACCAAGTATCGAGAGCAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCCCGAGTTGCTGCACTTTGTGGACCAAGTATCGAGAGCAGCTGATA 4560
 Qy 4561 GCCGAGTGACATCGGTGGAGGTTGCTTGGACAACTGCAATGACAGTGCAGCTGAGCCAG 4620
 Db 4561 GCCGAGTGACATCGGTGGAGGTTGCTTGGACAACTGCAATGACAGTGCAGCTGAGCCAG 4620
 Qy 4621 GAGCAGTACGAGAGGTTGCTGCTGAGAACACAGAGCCCGACAGATCGGAAGCTGTTTC 4680
 Db 4621 GAGCAGTACGAGAGGTTGCTGCTGAGAACACAGAGCCCGACAGATCGGAAGCTGTTTC 4680
 Qy 4681 AGCTTGAAGCAGTCTGCTGGACCGGAAGTCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Db 4681 AGCTTGAAGCAGTCTGCTGGACCGGAAGTCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Qy 4741 ACCCATCTCACTCATTTATGGAACCTCTGGAGAGGGGACGCAAAAGGGACTCTCTGCCA 4800
 Db 4741 ACCCATCTCACTCATTTATGGAACCTCTGGAGAGGGGACGCAAAAGGGACTCTCTGCCA 4800
 Qy 4801 CTCAGCAGTCAAGTATCAACACAGAGCCCTTGAACCTTGAAGTCTCTGGCTTTGGCTGACCC 4860
 Db 4801 CTCAGCAGTCAAGTATCAACACAGAGCCCTTGAACCTTGAAGTCTCTGGCTTTGGCTGACCC 4860
 Qy 4861 TTTCTTGGGTCTCAGTTTCTTCTGCAAAACAGTTGCCATCTGTTGCTTCCAGCA 4920
 Db 4861 TTTCTTGGGTCTCAGTTTCTTCTGCAAAACAGTTGCCATCTGTTTGGCTTCCAGCA 4920
 Qy 4921 CTAAAGTAAATGGAACCTTGTATGATGCTTGTCTGGCATTTATGTTCCATCCAGGGATG 4980
 Db 4921 CTAAAGTAAATGGAACCTTGTATGATGCTTGTCTGGCATTTATGTTCCATCCAGGGATG 4980
 Qy 4981 CCACAGGGGGCCCCAGTCCAGTGGCCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Db 4981 CCACAGGGGGCCCCAGTCCAGTGGCCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Qy 5041 GGCAGAGCCCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCAAGCCTAGA 5100
 Db 5041 GGCAGAGCCCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCAAGCCTAGA 5100
 Qy 5101 GCCCTCCGATCCCATCCAGGGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTC 5160
 Db 5101 GCCCTCCGATCCCATCCAGGGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTC 5160
 Qy 5161 TGGCTGTCTCAGGGTCCAGCCCAAAATTTGGGGTTTCAAGCTGGGAGCCACGTGGATTC 5220
 Db 5161 TGGCTGTCTCAGGGTCCAGCCCAAAATTTGGGGTTTCAAGCTGGGAGCCACGTGGATTC 5220
 Qy 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCAACAGAGTAAAGTGAAGGAAGTTT 5280
 Db 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCAACAGAGTAAAGTGAAGGAAGTTT 5280
 Qy 5281 ATTCAAAAAATAAGGAGTATCAGAGCTTTTGAATTTCTAGCAGGCTTTCCAGTT 5340
 Db 5281 ATTCAAAAAATAAGGAGTATCAGAGCTTTTGAATTTCTAGCAGGCTTTCCAGTT 5340
 Qy 5341 TTTTACAGAAAAACCCCTATAAATTTAAAAATTTTCTTAAATTTTAAAGTAAAAAAT 5400
 Db 5341 TTTTACAGAAAAACCCCTATAAATTTAAAAATTTTCTTAAATTTTAAAGTAAAAAAT 5400
 Qy 5401 ACAAAAAAGAAAAATGAAAAATAAGGAATGAAGTGTACCTAC 5444
 Db 5401 ACAAAAAAGAAAAATGAAAAATAAGGAATGAAGTGTACCTAC 5444

1896 TCGGACACAGCTCTGCAGAACTCTATTCCTCTTTTGGAGCAGGACGTTGGGTAGAGGT 1955
 1533 TCGGACACAGCTCTGCAGAACTCTATTCCTCTTTTGGAGCAGGACGTTGGGTAGAGGT 1592
 1956 CTTGGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTACAGATGAAG 2015
 1593 CTTGGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTACAGATGAAG 1652
 2016 GCAAGCAATTAGAGCCTTTAGGTTGGTCAAAATCAACAAAGAGCTCTGGCCCTGTGCT 2075
 1653 GCAAGCAATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGCCCTGTGCT 1712
 2076 TGTGGCCTGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTACAGATGAAG 1772
 1713 TGTGGCCTGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTACAGATGAAG 1772
 2136 AAAATCACACTGACTTCCAGACCAACCAACCTCTCTCTACATTAACCTTTGCCAGGC 2195
 1773 AAAATCACACTGACTTCCAGACCAACCAACCTCTCTCTACATTAACCTTTGCCAGGC 1832
 2196 TCTCCAGCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGG 2255
 1833 TCTCCAGCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGG 1892
 2256 CATCTGGCAAAAAGACCTTTTCACTCCAGATGACCTCAGAGACCTCTGCTCTCTGGCTGCTGAGG 2315
 1893 CATCTGGCAAAAAGACCTTTTCACTCCAGATGACCTCAGAGACCTCTGCTCTCTGGCTGCTGAGG 1952
 2316 GGCATCATCTCCACCTCTTGAAGATGGGTATTTCAAGACCAACCTCTCTGCTGAG 2375
 1953 GGCATCATCTCCACCTCTTGAAGATGGGTATTTCAAGACCAACCTCTCTGCTGAG 2012
 2376 CTACAGCTTCACTCAGCTCTGTTTCCAAGATCTTTTCAAGACCAACCTCTCTGCTGAG 2435
 2013 CTACAGCTTCACTCAGCTCTGTTTCCAAGATCTTTTCAAGACCAACCTCTCTGCTGAG 2072
 2436 GATGAGAGGGGAGAGAGTAAATCTAAATGTCATAGATTTGGAAGACGCTAGA 2495
 2073 GATGAGAGGGGAGAGAGTAAATCTAAATGTCATAGATTTGGAAGACGCTAGA 2132
 2496 AGCATATGAAATACATGAGCTGTTGGGACATCAACACACGTTTCTATTTGGGCTGTT 2555
 2133 AGCATATGAAATACATGAGCTGTTGGGACATCAACACACGTTTCTATTTGGGCTGTT 2192
 2556 AAGTATGAGGGGAGAGAGATGAGAAACATCTTTCACTCCGCTGCTCTCAGGGGAG 2615
 2193 AAGTATGAGGGGAGAGAGATGAGAAACATCTTTCACTCCGCTGCTCTCAGGGGAG 2252
 2616 GAACCTGATGAGTGGTCCCGTCCCTGAGCTGCTGCTGAGCAGCAGCTCTCTGGAGTC 2675
 2253 GAACCTGATGAGTGGTCCCGTCCCTGAGCTGCTGCTGAGCAGCAGCTCTCTGGAGTC 2312
 2676 CTTCCACTGCTTTGACGAGCTCGGAACAAAACGTTCTTGACACAGTGAATGGCCATTT 2735
 2313 CTTCCACTGCTTTGACGAGCTCGGAACAAAACGTTCTTGACACAGTGAATGGCCATTT 2372
 2736 CGAAGAAATGGGCAATGTTGAGAAACACAGATGAGCTCTTAGTGTGACATTTCTGCAT 2795
 2373 CGAAGAAATGGGCAATGTTGAGAAACACAGATGAGCTCTTAGTGTGACATTTCTGCAT 2432
 2796 TAAATTCAGCCGCCAGCTTCAAGAGCTTCAAGTATTGAGGACGAGCAGCAGATCAAC 2855
 2433 TAAATTCAGCCGCCAGCTTCAAGAGCTTCAAGTATTGAGGACGAGCAGCAGATCAAC 2492
 2856 ATGGAGCCCCACCATGGTATGCTGTTAGGTGGGTCCAGTCCAGATGCTTATTTGGCA 2915
 2493 ATGGAGCCCCACCATGGTATGCTGTTAGGTGGGTCCAGTCCAGATGCTTATTTGGCA 2552
 2916 GATTCTCTCTCGCTCCTCAAGGTCCACAGAAACCTGAGAGGAGCTTGACCTTAAGTGA 2975
 2553 GATTCTCTCTCGCTCCTCAAGGTCCACAGAAACCTGAGAGGAGCTTGACCTTAAGTGA 2612

2976 CTGCTGAGGACCACTCTGAGTGAAGAGTCTTTGTAAGACCTGAGACGCTCTGCTGCTGCT 3035
 2613 CTGCTGAGCCACTCTGAGTGAAGAGTCTTTGTAAGACCTGAGACGCTCTGCTGCTGCT 2672
 3036 CTTGGAGACCTTGGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3095
 2673 CTTGGAGACCTTGGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2732
 3096 CTTTGGGCTGAGAGCAACAGACCTCTGACGAGCTGAGACCTGAGCTTCAATGTGCTCAC 3155
 2733 CTTTGGGCTGAGAGCAACAGACCTCTGACGAGCTGAGACCTGAGCTTCAATGTGCTCAT 2792
 3156 GATGCTGAGAGCAACACCTTTTGCAGAGACTGAGACGCGAGCTCAAGCTACAGCG 3215
 2793 GATGCTGAGAGCAACACCTTTTGCAGAGACTGAGACGCGAGCTCAAGCTACAGCG 2852
 3216 ACTGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3275
 2853 ACTGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2912
 3276 TAGTGCAGCCCTGAGAGCTGAGCTGCTGAGCAGAAACCTTGGATGAGCTTGG 3335
 2913 TAGTGCAGCCCTGAGAGCTGAGCTGCTGAGCAGAAACCTTGGATGAGCTTGG 2972
 3336 CTTGCGACTGCTGAGAGGCTCAGGCATCTCTGCTGCAAACTCATACGCTGGGCT 3395
 2973 CTTGCGACTGCTGAGAGGCTCAGGCATCTCTGCTGCAAACTCATACGCTGGGCT 3032
 3396 GATGAGACCACTCTGAGTGAAGAGCTGAGCAGGAGCTGAGGCTTGGAGCAGGAGAA 3455
 3033 GATGAGACCACTCTGAGTGAAGAGCTGAGCAGGAGCTGAGGCTTGGAGCAGGAGAA 3092
 3456 ACTGAGCTGCTCATCTTTCAGCAGAGGAAACCTGAGTGTGATGACCTTCTGAGGGCT 3515
 3093 GCTGAGCTGCTCATCTTTCAGCAGAGGAAACCTGAGTGTGATGACCTTCTGAGGGCT 3152
 3516 GATGAGGAGAGTGAATAGCACAATCTCACTCAAGCGGAGAGACTCGGATCAGA 3575
 3153 GATGAGGAGAGTGAATAGCACAATCTCACTCAAGCGGAGAGACTCGGATCAGA 3212
 3576 GAGGCGGCTTCCATGTTGCTCAGGCTAACTCAAACTCTGAGCAGTGAAGATCTT 3635
 3213 GAGGCGGCTTCCATGTTGCTCAGGCTAACTCAAACTCTGAGCAGTGAAGATCTT 3272
 3636 CCAATTCCTGAGATTGAGAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTCTTGTG 3695
 3273 CCAATTCCTGAGATTGAGAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTCTTGTG 3332
 3696 CTTGCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCTTTGGGAGCTGAGATGA 3755
 3333 CTTGCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCTTTGGGAGCTGAGATGA 3392
 3756 CTTTGGGGCCCAAGGGGCTTGGCTGCTGAGGTAGTGTACAAAGAAAGAACTTGTGA 3815
 3393 CTTTGGGGCCCAAGGGGCTTGGCTGCTGAGGTAGTGTACAAAGAAAGAACTTGTGA 3452
 3816 CCGAGTTCACTTCCCTGAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTGT 3875
 3453 CCGAGTTCACTTCCCTGAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTGT 3512
 3876 GATGAGAGAGCGGTGAGCCTGAGATTTGAATTTCTGTTGGGACCAAGTTCTTGGGTGA 3935
 3513 GGTGAGAGAGCGGTGAGCCTGAGATTTGAATTTCTGTTGGGACCAAGTTCTTGGGTGA 3572
 3936 GATCAACCCACAGCAGCTGAGTGGGAGGCTCTGCTGAGCAGTCAAGGCTGAGCC 3995
 3573 GATCAACCCACAGCAGCTGAGTGGGAGGCTCTGCTGAGCAGTCAAGGCTGAGCC 3632
 3996 TGGAGCTGTTGAAGCTGACCTCCCTCATCTTTGTTGGCTCTCCAAAGGGGCGCATGTGA 4055
 3633 TGGAGCTGTTGAAGCTGACCTCCCTCATCTTTGTTGGCTCTCCAAAGGGGCGCATGTGA 3692
 4056 CACATCCCTGTTCCAAATGGCCCACTTTTAAAGAGAGGGGATGCTCTTGGAGAACCCAGC 4115

Db 3693 CACATCCCTGTTCCAAAGTGGCCCACTTTAAAGAGGAGGGATGCTCCTGGAGAGCCAGC 3752
 QY 4116 CAGGGTGGAGCTGCATCACATAGTCTTGGAAGACCCAGCTTCTCCCTCTGGAGTCTCT 4175
 Db 3753 CAGGGTGGAGCTGCATCACATAGTCTTGGAAGACCCAGCTTCTCCCTCTGGAGTCTCT 3812
 QY 4176 CTTGAAAATGATCCATAAATGCCCTCGGCTTCATTCCTGTCACCTCTGTGGTGTGCTTTA 4235
 Db 3813 CTTGAAAATGATCCATAAATGCCCTCGGCTTCATTCCTGTCACCTCTGTGGTGTGCTTTA 3872
 QY 4236 CCACGGGTCCATCTGAGAGAGTCACTTCCACCTCTACCTGATCCCAAGTCACTGCTC 4295
 Db 3873 CCACGGGTCCATCTGAGAGAGTCACTTCCACCTCTACCTGATCCCAAGTCACTGCTC 3932
 QY 4296 CATTCCGAAGGAATCGAGCTCTGCTATCGAAGCCCTGGAGAGACCCAGCTGTTCCTCGGA 4355
 Db 3933 CATTCCGAAGGAATCGAGCTCTGCTATCGAAGCCCTGGAGAGACCCAGCTGTTCCTCGGA 3992
 QY 4356 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTCAGGCTGCAAGTGAAGACCAAGAGATCA 4415
 Db 3993 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTCAGGCTGCAAGTGAAGACCAAGAGATCA 4052
 QY 4416 GACTCTGGTGTGGAGGCTTGGTGAACCCAGGAGATCTCATGCTGCAACTCTGAT 4475
 Db 4053 GACTCTGGTGTGGAGGCTTGGTGAACCCAGGAGATCTCATGCTGCAACTCTGAT 4112
 QY 4476 CCTCTCAGCCCGCATAGCCGTACTTCACTCTGGATGCCCGCAGTGTGCTGCACTTGT 4535
 Db 4113 CCTCTCAGCCCGCATAGCCGTACTTCACTCTGGATGCCCGCAGTGTGCTGCACTTGT 4172
 QY 4536 GGAACAGTATCGAGAGCAGCTGATAGCCGAGTACATCGGTGGAGGTTGTCTTGGACAA 4595
 Db 4173 GGAACAGTATCGAGAGCAGCTGATAGCCGAGTACATCGGTGGAGGTTGTCTTGGACAA 4232
 QY 4596 ACTGATGAGCAGGTGCTGAGCAGAGCAGTACGAGAGGCTGCTGGCTGAGAACACAGAG 4655
 Db 4233 ACTGATGAGCAGGTGCTGAGCAGAGCAGTACGAGAGGCTGCTGGCTGAGAACACAGAG 4292
 QY 4656 GCGCCAGCAGATCGGAAGCTGTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAAGA 4715
 Db 4293 GCGCCAGCAGATCGGAAGCTGTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAAGA 4352
 QY 4716 TGACTCTTACCAAGCCCTGAAGAGAGCCATCTCCTCATTTATGAACTCTGGGAGAA 4775
 Db 4353 TGACTCTTACCAAGCCCTGAAGAGAGCCATCTCCTCATTTATGAACTCTGGGAGAA 4412
 QY 4776 GGGCAGAAAAGGGACTCTGCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACC 4835
 Db 4413 GGGCAGAAAAGGGACTCTGCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACC 4472
 QY 4836 CTTGAGTCCCTGCTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTCTGCAAAACAG 4895
 Db 4473 CTTGAGTCCCTGCTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTCTGCAAAACAG 4532
 QY 4896 TTGCCATCTGGTTGGCTTCCAGCAGCTTAAAGTAAATGGAATTTGATGATGCTTCTGCTG 4955
 Db 4533 TTGCCATCTGGTTGGCTTCCAGCAGCTTAAAGTAAATGGAATTTGATGATGCTTCTGCTG 4592
 QY 4956 GCATTATGTTGTCATGCCAGGATGCAACAGGGGGCCCAAGTCCAGTGGCTTAAACAGCA 5015
 Db 4593 GCATTATGTTGTCATGCCAGGATGCAACAGGGGGCCCAAGTCCAGTGGCTTAAACAGCA 4652
 QY 5016 TCTCAGGAGTGTCCATCTGGAGTGGCAAGCCCTGCAAGCTCATAGAGCTCATCT 5075
 Db 4653 TCTCAGGAGTGTCCATCTGGAGTGGCAAGCCCTGCAAGCTCATAGAGCTCATCT 4712
 QY 5076 GTTGGCCACAGCAGCCAGGCTTAGAGCCCTCCGATCCCATCCAGCGCAAGAGCAATA 5135
 Db 4713 GTTGGCCACAGCAGCCAGGCTTAGAGCCCTCCGATCCCATCCAGCGCAAGAGCAATA 4772
 QY 5136 GGAGGACATGGAACCATTTGCTCTGGCTGTGCTCAGGGGTGAGCCCAAAATTTGGGT 5195

Db 4773 GGAGGACATGGAACCATTTGCTCTGGCTGTGTGCAGGGTGAGCCCAAAATTTGGGT 4832
 QY 5196 TCAGCGTGGAGGCCACCGTGGATTTCTGGCTTTGTACAGGAAGATCTACAAGAGCAAGCC 5255
 Db 4833 TCAGCGTGGAGGCCACCGTGGATTTCTGGCTTTGTACAGGAAGATCTACAAGAGCAAGCC 4892
 QY 5256 AACAGAGTAAAGTGAAGAGTATTATTCAGAAAATAAAGAGTATCACAGCTCTTTAG 5315
 Db 4893 AACAGAGTAAAGTGAAGAGTATTATTCAGAAAATAAAGAGTATCACAGCTCTTTAG 4952
 QY 5316 AATTTGTCTAGCAGGCTTTCCAGTTTTACCAAGAAACCCCTATATAATTAATAATTTT 5375
 Db 4953 AATTTGTCTAGCAGGCTTTCCAGTTTTACCAAGAAACCCCTATATAATTAATAATTTT 5012
 QY 5376 ACTTAAATTAAGAAATTAATAATTAATAATTAATAATTAATAATTAAGAAATAAGAA 5435
 Db 5013 ACTTAAATTAAGAAATTAATAATTAATAATTAATAATTAATAATTAAGAAATAAGAA 5072
 QY 5436 GTTA 5439
 Db 5073 GTTA 5076

RESULT 5
 AAD12951
 ID AAD12951 standard; cDNA; 5122 BP.
 XX
 AC AAD12951;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human G-protein coupled receptor-8 (GCRC-8) cDNA.
 KW Human; G-protein coupled receptor-8; GCRC-8; cytostatic; hepatotropic;
 KW viricide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective;
 KW nootropic; cerebroprotective; hypotensive; tranquilizer; vulnerary;
 KW ophthalmological; cell proliferative disorder; actinic keratosis;
 KW anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
 KW psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease;
 KW Huntington's disease; Parkinson's disease; cardiovascular disorder;
 KW epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia;
 KW anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder;
 KW Addison's disease; Crohn's disease; acquired immune deficiency syndrome;
 KW AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity;
 KW osteoporosis; transgenic animal; gene therapy; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 489..4910
 FT /*tag= a
 FT /product= "Human GCRC-8 protein"
 XX
 PN WO200157085-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 01-FEB-2001; 2001WO-US03455.
 XX
 PR 02-FEB-2000; 2000US-0180093.
 PR 11-FEB-2000; 2000US-0182045.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Baughn MR, Au-Young J, Yue H;
 XX WPI; 2001-488869/53.
 DR P-PSDB; AAE06758.
 DR
 XX Novel isolated human G-protein coupled receptor useful for diagnosing,
 PT preventing and treating cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic
 PT disorders -

XX The present sequence is human G-protein coupled receptor-8 (GRCR-8)
PS cDNA. The present invention relates to GRCR protein and nucleic acids
CC encoding them. GRCR protein, its agonist or antagonist are useful for
CC treating diseases or conditions associated with decreased expression
CC or overexpression of functional GRCR in a patient, where the disorder
CC is selected from cell proliferative disorders such as actinic keratosis,
CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and
CC cancer, neurological disorders such as epilepsy, stroke, Alzheimer's
CC disease, Huntington's disease, Parkinson's disease, cardiovascular
CC disorders such as hypertension, vasculitis, varicose veins, gastro-
CC intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea,
CC pancreatitis, autoimmune/inflammatory disorders such as acquired
CC immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease,
CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC infections, trauma and metabolic disorders such as diabetes, obesity,
CC osteoporosis. GRCR proteins and their cDNAs are used to assess the
CC effects of exogenous compounds on the expression of GRCR sequences.
CC GRCR cDNA is useful to create knock in humanised animals (pigs) or
CC transgenic animals (mice or rats) to model human disease, for
CC therapeutic or diagnostic purposes, for somatic or germline gene
CC therapy, to generate hybridisation probes useful in mapping the
CC naturally occurring genomic sequence, and in molecular biological
CC techniques.
XX
SQ Sequence 5122 BP; 1184 A; 1460 C; 1401 G; 1077 T; 0 other;

Query Match 87.5%; Score 4765.6; DB 22; Length 5122;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 4933; Conservative 0; Mismatches 24; Indels 133; Gaps 2;

QY 35 GAGCCAGACGCGGGCTCCACTGCGGTCTGAAAGCCCATTCCTCTCTGGGCTC 94
DB 1 GAGCCAGACGCGGGCTCCACTGCGGTCTGAAAGCCCATTCCTCTCTGGGCTC 60

QY 95 CTCGCCACCCACTCTTCAGCCTTTCAGCTTTCAGGCTTTCAGGAGTCCAGGACC 154
DB 61 CTCGCCACCCACTCTTCAGCCTTTCAGCTTTCAGGAGTTCAGGAGTCCAGGACC 120

QY 155 CAGGAGAGGAGAAATCTGAGGAACACAGAACAGTGGGCTTGCACACCCCATCTCC 214
DB 121 CAGGAGAGGAGAAATCTGAGGAACACAGAACAGTGGGCTTGCACACCCCATCTCC 180

QY 215 GTCACACATCTCCCTCACCCTCACCCTCCCTGCGCTGCGCTGCGCTGCGCTGCG 274
DB 181 GTCACACATCTCCCTCACCCTCACCCTCCCTGCGCTGCGCTGCGCTGCGCTGCG 240

QY 275 CTCCTTATCAGCTGACTTCTTCAGTGTCTTGAGGCGCCCTCTGCGCTCCTCCCTCC 334
DB 241 CTCCTTATCAGCTGACTTCTTCAGTGTCTTGAGGCGCCCTCTGCGCTCCTCCCTCC 300

QY 335 GGCTTTTCTTACACCTCCCTCTATCGGCTCTATCTGTAGTGCCCTGGGATTTATAA 394
DB 301 GGCTTTTCTTACACCTCCCTCTATCGGCTCTATCTGTAGTGCCCTGGGATTTATAA 360

QY 395 AACTGGGTTCCGATCTGAAATAGAGACGGTAAAGCCCAAGCAAGACAGCACTGTT 454
DB 361 AACTGGGTTCCGATCTGAAATAGAGACGGTAAAGCCCAAGCAAGACAGCACTGTT 420

QY 455 CTCCTGCTCCTGATACCTCACCCTCGGAAACATCCCGACACCTCTTTACTCCG 514
DB 421 CTCCTGCTCCTGATACCTCACCCTCGGAAACATCCCGACACCTCTTTACTCCG 480

QY 515 GGAACAGATGGTGGCGAGCTGGGCGCGCTGCGCTGTTACTTTGGAGTTCTTGAAGA 574
DB 481 GGAACAGATGGTGGCGAGCTGGGCGCGCTGCGCTGTTACTTTGGAGTTCTTGAAGA 540

QY 575 AGGAGAGCTGAAGAGGTTCCAGCTTCTGCTGCGCAATAAAGGCACTCCAGAGCTCTT 634
DB 541 AGGAGAGCTGAAGAGGTTCCAGCTTCTGCTGCGCAATAAAGGCACTCCAGAGCTCTT 600

QY 635 CGGTGAGACACCCGCTCAGCCAGAGAAAGACAGTGGCATGGAGTGGCTCGTACCTGG 694
DB 601 CGGTGAGACACCCGCTCAGCCAGAGAAAGACAGTGGCATGGAGTGGCTCGTACCTGG 660

QY 695 TGGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGG 754
DB 661 TGGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGG 720

QY 755 GGCTGAGTCACTGTGCGCCCAAGCCAGAGAGGCGCAGGCACTCTCCCTCATTTCCCT 814
DB 721 GGCTGAGTCACTGTGCGCCCAAGCCAGAGAGGCGCAGGCACTCTCCCTCATTTCCCT 780

QY 815 ACAGCCCAAGTAAACCCCACTGGGCTTCCAGCAACCCCACTTCCACCGCAGTGCTAA 874
DB 781 ACAGCCCAAGTAAACCCCACTGGGCTTCCAGCAACCCCACTTCCACCGCAGTGCTAA 840

QY 875 TGCCCTGATCCATGAATTCGCGGGGGTGCACCCAGGCTCAGAGAGAGGTTTTGA 934
DB 841 TGCCCTGATCCATGAATTCGCGGGGGTGCACCCAGGCTCAGAGAGAGGTTTTGA 900

QY 935 GACAGCTCCTTGACACATCTGAGCGCGCTGAGAGAAATCTTGCCTCACTCTCTTACC 994
DB 901 GACAGCTCCTTGACACATCTGAGCGCGCTGAGAGAAATCTTGCCTCACTCTTACC 960

QY 995 AAGCTTTTCAAGCTTCCCGCAGACCATGAGTCTCCAAGCCAGGAGTCAACCAACGCCCA 1054
DB 961 AAGCTTTTCAAGCTTCCCGCAGACCATGAGTCTCCAAGCCAGGAGTCAACCAACGCCCA 1020

QY 1055 CATCCACAGCAGTGTGGGAGCTGGGGATCCCACTCAGCCCAAGCTTAGCACCAG 1114
DB 1021 CATCCACAGCAGTGTGGGAGCTGGGGATCCCACTCAGCCCAAGCTTAGCACCAG 1080

QY 1115 AGCAGGAGCTCTCTGGGACCAATGGGCTCTGGATGAAACGTGAGAAATTTACTACACAG 1174
DB 1081 AGCAGGAGCTCTCTGGGACCAATGGGCTCTGGATGAAACGTGAGAAATTTACTACACAG 1140

QY 1175 AAATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGAGGCGCCCATGGGAGCGGTG 1234
DB 1141 AAATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGAGGCGCCCATGGGAGCGGTG 1200

QY 1235 TAGGAACGCGCCCAAGCGGACACAGCCTTACAGCCCAACACACCCATGGGAGCCTT 1294
DB 1201 TAGGAACGCGCCCAAGCGGACACAGCCTTACAGCCCAACACACCCATGGGAGCCTT 1260

QY 1295 CTGTGAGAGAGAGCTCTGTTTCCATGCGCTTGGAAATGAGGATTTAAACCAAAAT 1354
DB 1261 CTGTGAGAGAGAGCTCTGTTTCCATGCGCTTGGAAATGAGGATTTAAACCAAAAT 1320

QY 1355 TCACACAGCTGCTACTTCTACAAAGACCTTCCAGAGCAAGATCCCTCTGTGCAAGA 1414
DB 1321 TCACACAGCTGCTACTTCTACAAAGACCTTCCAGAGCAAGATCCCTCTGTGCAAGA 1380

QY 1415 GAAGTGGCTGATTTATGTGAGAGAGAAATCAGAGCAATTTAATTTGAGATCAGAGCTTAT 1474
DB 1381 GAAGTGGCTGATTTATGTGAGAGAGAAATCAGAGCAATTTAATTTGAGATCAGAGCTTAT 1440

QY 1475 TTGGCCCAAGGCTGATACCAAGACCTTCCATAGTATCATCTGAGGAGGCTGTGGA 1534
DB 1441 TTGGCCCAAGGCTGATACCAAGACCTTCCATAGTATCATCTGAGGAGGCTGTGGA 1500

QY 1535 TTGGGAAGTCAACACTGGCCAGGAGGAGTGAAGAGCCCTGGGGAGAGGAGCCAGCTGTATG 1594
DB 1501 TTGGGAAGTCAACACTGGCCAGGAGGAGTGAAGAGCCCTGGGGAGAGGAGCCAGCTGTATG 1560

QY 1595 GGAACCGCTTCCAGCATGCTTTTACTTTCAGTGGAGAGCTGGCCCAAGCTTCAAGGTGG 1654
DB 1561 GGAACCGCTTCCAGCATGCTTTTACTTTCAGTGGAGAGCTGGCCCAAGCTTCAAGGTGG 1620

QY 1655 TGAGTCTCGCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGCTCCCATTTAGACAGA 1714
DB 1621 TGAGTCTCGCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGCTCCCATTTAGACAGA 1680

QY 1715 TCCTGCTTAGGCCAGAGCGGCTGCTTTTCATCTCTGATGGTGTAGATGAGCCAGGATGGG 1774

Db 1681 TCTCTGTAGGCGAGCGGCTCTTCTATCTCGATGGTGTAGATGAGCGAGATGGG 1740
 Qy 1775 TCTTCAGAGCGCGAGTCTGTAGCTCTGTCTGCTAGCTGAGCGAGCGAGCGAGCGGATG 1834
 Db 1741 TCTTCAGAGCGCGAGTCTGTAGCTCTGTCTGCTAGCTGAGCGAGCGAGCGGATG 1800
 Qy 1835 CACTGCTGGCAGTTTGTCTGGGAAAACTATATCTTCCGAGGCACTCTTCTGTGATCACGG 1894
 Db 1801 CACTGCTGGCAGTTTGTCTGGGAAAACTATATCTTCCGAGGCACTCTTCTGTGATCACGG 1860
 Qy 1895 CTCGACACACAGCTCTGCAGAACCTTCTCTTCTTGGAGGAGCGAGCTTGGGTAGAGG 1954
 Db 1861 CTCGACACACAGCTCTGCAGAACCTTCTCTTCTTCTTGGAGGAGCGAGCTTGGGTAGAGG 1920
 Qy 1955 TCTCTGGGTCTCTCAGTCTGCAGCGAGGAGGATATTTCTTACAGATATTTTACAGATGAAA 2014
 Db 1921 TCTCTGGGTCTCTCAGTCTGCAGCGAGGAGGATATTTCTTACAGATATTTTACAGATGAAA 1980
 Qy 2015 GGCAAGCAATTAGAGCTTTAGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCT 2074
 Db 1981 GGCAAGCAATTAGAGCTTTAGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCT 2040
 Qy 2075 TTGTCCCTGGGTCTCTGGCTGGCTGCACTTGGCTGATGAGCAGATGAAGCGGAAGG 2134
 Db 2041 TTGTCCCTGGGTCTCTGGCTGGCTGCACTTGGCTGATGAGCAGATGAAGCGGAAGG 2100
 Qy 2135 AAAAACTCACCTGACTTCCAAAGACACCAACACCTCTGTCTACATTACCTTGGCCAGG 2194
 Db 2101 AAAAACTCACCTGACTTCCAAAGACACCAACACCTCTGTCTACATTACCTTGGCCAGG 2160
 Qy 2195 CTCTCAAGCTCAGGCAATTGGGACCCAGCTCAGAGACCTCTGTCTCTCTGGCTGTGAGG 2254
 Db 2161 CTCTCAAGCTCAGGCAATTGGGACCCAGCTCAGAGACCTCTGTCTCTCTGGCTGTGAGG 2220
 Qy 2255 GCATCTGGCAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCAGTGGTGTAGATG 2314
 Db 2221 GCATCTGGCAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCAGTGGTGTAGATG 2280
 Qy 2315 GGGCCATCATCTCCACCTCTTGTAGATGGGTATTTCTTCAAGAGCACCCCATCTCTCTCA 2374
 Db 2281 GGGCCATCATCTCCACCTCTTGTAGATGGGTATTTCTTCAAGAGCACCCCATCTCTCTCA 2340
 Qy 2375 GCTACAGCTTCATTCACCTCTGTTTCCAAAGAGTCTTTTGCAGCAATGTCTATGTCTTGG 2434
 Db 2341 GCTACAGCTTCATTCACCTCTGTTTCCAAAGAGTCTTTTGCAGCAATGTCTATGTCTTGG 2400
 Qy 2435 AGGATGAGAGGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGGAAAAAGACGCTAG 2494
 Db 2401 AGGATGAGAGGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGGAAAAAGACGCTAG 2460
 Qy 2495 AAGCATATGGAATACATGGCCCTGTTTGGGGCATCAACACAGCTTCTTATTTGGGCCCTGT 2554
 Db 2461 AAGCATATGGAATACATGGCCCTGTTTGGGGCATCAACACAGCTTCTTATTTGGGCCCTGT 2520
 Qy 2555 TAAGTATGAGGGGAGAGAGATGGAGAAACATTTTCACTGCCGGCTGTCTCAGGGGA 2614
 Db 2521 TAAGTATGAGGGGAGAGAGATGGAGAAACATTTTCACTGCCGGCTGTCTCAGGGGA 2580
 Qy 2615 GGAACCTGATGAGTGGGTCCCGTCCCTGCAGTGTCTGTGAGCCACACTCTCTGGAGT 2674
 Db 2581 GGAACCTGATGAGTGGGTCCCGTCCCTGCAGTGTCTGTGAGCCACACTCTCTGGAGT 2640
 Qy 2675 CCTCCACTCTGTACGAGACTCGGAAACAAACGTTCTCTGACACAGGTGATGGGCCATT 2734
 Db 2641 CCTCCACTCTGTATGAGACTCGGAAACAAACGTTCTCTGACACAGGTGATGGGCCATT 2700
 Qy 2735 TCGAAGAAATGGSCATGTGTAGAAACAGACATGAGCTCTTAGTGTGCACTTTCTGCA 2794
 Db 2701 TCGAAGAAATGGSCATGTGTAGAAACAGACATGAGCTCTTAGTGTGCACTTTCTGCA 2760
 Qy 2795 TTAAATTCAGCGCCACGTTGAAGAGCTTTCAGTGTAGTGGGCGAGGACGACAGATCAA 2854

Db 2761 TTAAATTCAGCGCCACGTTGAAGAGCTTTCAGCTGATTTGAGGGCAGGACGACAGATCAA 2820
 Qy 2855 CATGGAGCCCCACCATGCTAGTCTCTGTTTTCAGTGGGTCCCACTCAGATCCCTATTGGC 2914
 Db 2821 CATGGAGCCCCACCATGCTAGTCTCTGTTTTCAGTGGGTCCCACTCAGATCCCTATTGGC 2880
 Qy 2915 AGATTCTTCTTCTCGTCTCAAGGTCCACAGAACTTGAAGAGCTGACCTTAAGTGAA 2974
 Db 2881 AGATTCTTCTTCTCGTCTCAAGGTCCACAGAACTTGAAGAGCTGACCTTAAGTGAA 2940
 Qy 2975 ACTCGCTGAGCACCTCTGCACTGAAGAGTCTTTTGAAGACCTTGAAGCGCTCTCGCTGCC 3034
 Db 2941 ACTCGCTGAGCACCTCTGCACTGAAGAGTCTTTTGAAGACCTTGAAGCGCTCTCGCTGCC 3000
 Qy 3035 TCTTGGAGACCTCTCGGTGTGGCTGTGGCTCTCACAGCTGAGGACTGCAAGGACCTTG 3094
 Db 3001 TCTTGGAGACCTCTCGGTGTGGCTGTGGCTCTCACAGCTGAGGACTGCAAGGACCTTG 3060
 Qy 3095 CTTTGGGCTGAGAGCCAAACAGACCTTGAACCGAGCTGGAGCTTGAGCTTCAATGTGCTCA 3154
 Db 3061 CTTTGGGCTGAGAGCCAAACAGACCTTGAACCGAGCTGGAGCTTGAGCTTCAATGTGCTCA 3120
 Qy 3155 CGGATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTTACAGC 3214
 Db 3121 CGGATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTTACAGC 3180
 Qy 3215 GACTGCACTGCTGAGCTGTGGCTCTCACGTCTGACTGTGCGCAGGACCTGGCCTCTGTGC 3274
 Db 3181 GACTGCACTGCTGAGCTGTGGCTCTCACGTCTGACTGTGCGCAGGACCTGGCCTCTGTGC 3240
 Qy 3275 TTAGTCCAGGCCAGCCCTGAAGAGCTAGACCTTGCAGCAGAAACAACTTGAATGACGTTG 3334
 Db 3241 TTAGTCCAGGCCAGCCCTGAAGAGCTAGACCTTGCAGCAGAAACAACTTGAATGACGTTG 3300
 Qy 3335 CGGTGCACTGCTCTGTCAGGGGCTCAGGCATCTCTGCTGCAAACTCATAGCCCTGGGCG 3394
 Db 3301 CGGTGCACTGCTCTGTCAGGGGCTCAGGCATCTCTGCTGCAAACTCATAGCCCTGGGCG 3360
 Qy 3395 TGGACACAGCAACTCTGAGTGTATGAGTGAAGGAGAACTGAGGCGCTTGAAGCAGGAGA 3454
 Db 3361 TGGACACAGCAACTCTGAGTGTATGAGTGAAGGAGAACTGAGGCGCTTGAAGCAGGAGA 3420
 Qy 3455 AACCTCAGCTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCC 3514
 Db 3421 AACCTCAGCTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCC 3480
 Qy 3515 TGGATACGGGAGAGATGATTAATAGCACATCTCTCACTCAAGCGGACAGACTCCGATCAG 3574
 Db 3481 TGGATACGGGAGAGATGATTAATAGCACATCTCTCACTCAAGCGGACAGACTCCGATCAG 3540
 Qy 3575 AGAGGGCGCTTCCCATGCTTGTCTCAGGCTAATCTCAAACTCTCTGACGCTGAGCAAGATCT 3634
 Db 3541 AGAGGGCGCTTCCCATGCTTGTCTCAGGCTAATCTCAAACTCTCTGACGCTGAGCAAGATCT 3600
 Qy 3635 TCCCAATTTGCTGAGATTGACAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGT 3694
 Db 3601 TCCCAATTTGCTGAGATTGACAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTGT 3660
 Qy 3695 GGTGCTCTTCTCTGCTCTCAAGGGACCTGCATACGAGCCTTTGGGACCTGACGATG 3754
 Db 3661 GGTGCTCTTCTCTGCTCTCAAGGGGACCTGCATACGAGCCTTTGGGACCTGACGATG 3720
 Qy 3755 ACTTCTGGGGCCCCACGGGCTGTGGCTTACTGAGGTAGTTGACAAAGAAAAAGAACTTGT 3814
 Db 3721 ACTTCTGGGGCCCCACGGGCTGTGGCTTACTGAGGTAGTTGACAAAGAAAAAGAACTTGT 3780
 Qy 3815 ACCGAGTTCACTTCCCTGTAGTGGCTCTCAACCGTGGCCCCAACCGGTCTCTGCTTTG 3874
 Db 3781 ACCGAGTTCACTTCCCTGTAGTGGCTCTCAACCGTGGCCCCAACCGGTCTCTGCTTTG 3840
 Qy 3875 TGAATGAGAGAGCGGTGACGTTGAGTGAATTTCTGTGTGTGGGACCAAGTTCTCTGGGTG 3934
 Db 3841 TGAATGAGAGAGCGGTGACGTTGAGTGAATTTCTGTGTGTGGGACCAAGTTCTCTGGGTG 3900

Db	1981	GGAAATACATGGCTGTTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGTTAAAGTGAT	2040
Qy	2563	GAGGGGAGAGAGATGAGAGAACTCTTTACTGCGGCTGTCTCAGGGGAGGAACCTG	2622
Db	2041	GAGGGGAGAGAGATGAGAGAACTCTTTCACTGCGGCTGTCTCAGGGGAGGAACCTG	2100
Qy	2623	ATGCACTGGTCCGCTCCCTGAGCTGCTGTCAGCCACACTCTCTGGAGTCCCTCCAC	2682
Db	2101	ATGCACTGGTCCGCTCCCTGAGCTGCTGTCAGCCACACTCTCTGGAGTCCCTCCAC	2160
Qy	2683	TGCTTTACGAGCTCGGAACAAACGTTCTTGACAAAGTATGATGCCCATTTGGAAGAA	2742
Db	2161	TGCTTTACGAGCTCGGAACAAACGTTCTTGACAAAGTATGATGCCCATTTGGAAGAA	2220
Qy	2743	ATGGGATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGACCTTTCTGCAATTAATTC	2802
Db	2221	ATGGGATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGACCTTTCTGCAATTAATTC	2280
Qy	2803	AGCGCCACGTGAAGAGCTTCAGCTGATTGAGGGCAGGCACACAGATCAACATGGAGC	2862
Db	2281	AGCGCCACGTGAAGAGCTTCAGCTGATTGAGGGCAGGCACACAGATCAACATGGAGC	2340
Qy	2863	CCACCATGTGTCTGTTTCAAGTGGGTCCAGTCAAGATGCCCTATTGGCAGATTTCT	2922
Db	2341	CCACCATGTGTCTGTTTCAAGTGGGTCCAGTCAAGATGCCCTATTGGCAGATTTCT	2400
Qy	2923	TTCTCGTCTCAAGTGTACAGAAACCTGAAGAGCTGGACTGAAGTGGAAATCTCGTG	2982
Db	2401	TTCTCGTCTCAAGTGTACAGAAACCTGAAGAGCTGGACTGAAGTGGAAATCTCGTG	2460
Qy	2983	AGCCACTCTGAGTGAAGAGTCTTTGTAAGACCTGAGAGCCCTCGCTGCTCTGGAG	3042
Db	2461	AGCCACTCTGAGTGAAGAGTCTTTGTAAGACCTGAGAGCCCTCGCTGCTCTGGAG	2520
Qy	3043	ACCCTCGGTGTGGCTGGCTTCAGCTCAGAGTGGAGCTGCAAGAGACCTTGCTTTGGG	3102
Db	2521	ACCCTCGGTGTGGCTGGCTTCAGCTCAGAGTGGAGCTGCAAGAGACCTTGCTTTGGG	2580
Qy	3103	CTGAGAGCAACAGACCTTGACCGAGCTGGACTGAGCTTCAATGTGTCTCAGATGCT	3162
Db	2581	CTGAGAGCAACAGACCTTGACCGAGCTGGACTGAGCTTCAATGTGTCTCAGATGCT	2640
Qy	3163	GGAGCAACACCTTTGCCAGAGCTGACAGCCGAGCTGCAAGCTACAGAGCTGCAG	3222
Db	2641	GGAGCAACACCTTTGCCAGAGCTGACAGCCGAGCTGCAAGCTACAGAGCTGCAG	2700
Qy	3223	CTGGTCAGCTGTGGCTTCAGCTGCTGCTGAGGACCTGGCTCTCTGCTTAGTGCC	3282
Db	2701	CTGGTCAGCTGTGGCTTCAGCTGCTGCTGAGGACCTGGCTCTCTGCTTAGTGCC	2760
Qy	3283	AGCCGAGCTGAGAGCTAGACTGCGAGAGAAACACTGGATGACTGTTGGCTGGCA	3342
Db	2761	AGCCGAGCTGAGAGCTAGACTGCGAGAGAAACACTGGATGACTGTTGGCTGGCA	2820
Qy	3343	CTGCTCTGTGAGGGCTCAGGATCTGCTGCAACTCATACGCTGGGCTGGACAG	3402
Db	2821	CTGCTCTGTGAGGGCTCAGGATCTGCTGCAACTCATACGCTGGGCTGGACAG	2880
Qy	3403	ACAACCTCTGAGTATGAGATGAGGAGGAACTGAGGGCTCTGAGCAGAGAGAACTCAG	3462
Db	2881	ACAACCTCTGAGTATGAGATGAGGAGGAACTGAGGGCTCTGAGCAGAGAGAACTCAG	2940
Qy	3463	CTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAG	3522
Db	2941	CTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAG	3000
Qy	3523	GGAGATGAGTAATAGCATCTCTCACTCAGCGGCAGAGCTCGGATCAGAGGGGCG	3582
Db	3001	GGAGATGAGTAATAGCATCTCTCACTCAGCGGCAGAGCTCGGATCAGAGGGGCG	3060
Qy	3583	GCTTCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAGTGAAGATCTTCCCAATT	3642

3061	GCTTCCATGTTGCTCAGGCTAAATCTCAAACTCTCGGACGTGAGCAAGATCTTCCCAATT	3120
3643	GCTGAGATTGAGAGGAAGCTCCCCAGAGTAGTACCGGTGGAACTCTTGTGCGTGCT	3702
3121	GCTGAGATTGAGAGGAAGCTCCCCAGAGTAGTACCGGTGGAACTCTTGTGCGTGCT	3180
3703	TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTTGGGACTGACGATGCTTCTGG	3762
3181	TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTTGGGACTGACGATGCTTCTGG	3240
3763	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGAAAAGAAAGAACTTGTACCGAGTT	3822
3241	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGAAAAGAAAGAACTTGTACCGAGTT	3300
3823	CACTTTCCCTGTAGCTGGCTCTACCGTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
3301	CACTTTCCCTGTAGCTGGCTCTACCGTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3360
3883	GAAGCGGTACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3942
3361	GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3420
3943	CCACAGCAGCTGATGGTGGCAGGGCTCTGTCTGACATCAAGGCTGAGCTGGAGCT	4002
3421	CCACAGCAGCTGATGGTGGCAGGGCTCTGTCTGACATCAAGGCTGAGCTGGAGCT	3480
4003	GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGACACATCC	4062
3481	GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGACACATCC	3540
4063	CTGTTTCCAAATGGGCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGGGGTG	4122
3541	CTGTTTCCAAATGGGCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGGGGTG	3600
4123	GAGCTGCATCATATGTTCTGGAAAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAA	4182
3601	GAGCTGCATCATATGTTCTGGAAAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAA	3660
4183	ATGATCCATATGCTCGGCTTCATTTCCCTGCTCACTCTGCTGTTGTTTACCAACGC	4242
3661	ATGATCCATATGCTCGGCTTCATTTCCCTGCTCACTCTGCTGTTTACCAACGC	3720
4243	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCG	4302
3721	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCG	3780
4303	-----	4302
3781	AAGGCCATAGATGATCTAGAAATGAAATTCAGTTTGTGCGAATCCACAAGCCACCCCG	3840
4303	-----	4302
3841	CTGACCCCACTTTATATGGGCTGTGTTTACTGTGTCTGGTCTGGTTTACGGGATGCTG	3900
4303	-----	4302
3901	GAATACTCCCCAGGAAGTCTGAGCTCTGCTATCGAAGCTCTGAGAGAGACAGGTTTC	3960
4351	TCGGAGTTCTAGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGAGAAA	4410
3961	TCGGAGTTCTAGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGAGAAA	4020
4411	GATGAGACTCTGGTGTGGAGGCTTGTGAAACAGAGAGATCTCATGCTGCAACTACT	4470
4021	GATGAGACTCTGGTGTGGAGGCTTGTGAAACAGAGAGATCTCATGCTGCAACTACT	4080
4471	CTGATCCCTCCAGCCCGCATAGCGTACTTCACTCTGGATGCCCGCAGTTGCTGCAC	4530
4081	CTGATCCCTCCAGCCCGCATAGCGTACTTCACTCTGGATGCCCGCAGTTGCTGCAC	4140
4531	TTTGTGGAACAGTATCGAGAGCAGCTGATAGCCGAGTGAATCGGTGGAGGTTCTTTG	4590
4141	TTTGTGGAACAGTATCGAGAGCAGCTGATAGCCGAGTGAATCGGTGGAGGTTCTTTG	4200

QY	4591	GACAAACTGCATGGACAGGTCTGTAGCCAGGAGCAGTACGAGAGGGTGCTGGCTGAGAAC	4650
Db	4201	GACAAACTGCATGGACAGGTCTGTAGCCAGGAGCAGTACGAGAGGGTGCTGGCTGAGAAC	4260
QY	4651	ACGAGGCCACGACGATGCGGAAAGCTTTTCAGCTTGAGCCAGTCTCGGACCGGAAGTGC	4710
Db	4261	ACGAGGCCACGACGATGCGGAAAGCTTTTCAGCTTGAGCCAGTCTCGGACCGGAAGTGC	4320
QY	4711	AAAGATGGACTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATTATGGAACTCTGG	4770
Db	4321	AAAGATGGACTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATTATGGAACTCTGG	4380
QY	4771	GAGAAGGGCAGCAAAAAGGGACTCTCTGCCACTCAGCAGCTGA	4812
Db	4381	GAGAAGGGCAGCAAAAAGGGACTCTCTGCCACTCAGCAGCTGA	4422
RESULT 7			
AAD02761			
ID	AAD02761 standard; cDNA; 4194 BP.		
XX			
AC	AAD02761;		
XX			
DT	31-MAY-2001 (first entry)		
XX			
DE	Human NB-ARC and CARD containing protein (NAC) gamma isoform cDNA.		
XX			
KW	Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;		
KW	caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;		
KW	cysteine aspartyl protease; apoptosis; cytokine production;		
KW	cytokine receptor signalling; therapy; inflammatory disorder; sepsis;		
KW	fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FD	1..4194		
CDS	/*tag= a		
FT	/product= "Human NB-ARC and CARD containing protein (NAC)		
FT	gamma isoform"		
FT	1..2868		
FT	misc_feature		
FT	/*tag= b		
FT	/note= "Corresponds to 1-2868 residues of human NAC		
FT	beta isoform (AAD02760)"		
FT	2869..3687		
FT	misc_feature		
FT	/*tag= c		
FT	/note= "Corresponds to 2962-3780 residues of human NAC		
FT	beta isoform (AAD02760)"		
FT	3688..4194		
FT	misc_feature		
FT	/*tag= d		
FT	/note= "Corresponds to 3916-4422 residues of human NAC		
FT	beta isoform (AAD02760)"		
XX			
PN	WO200116170-A2.		
XX			
PD	08-MAR-2001.		
XX			
PF	01-SEP-2000; 2000WO-US24152.		
XX			
PR	01-SEP-1999; 99US-0388221.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Reed JC;		
XX			
WP	WPI; 2001-183258/18.		
DR	P-PSDB; AAY72670.		
XX			
PT	Novel nucleic acid encoding NB-ARC and caspase associated recruitment		
PT	domains, used to produce polypeptides for screening for modulators of		
PT	apoptosis -		
XX			

PS Claim 4; Page 138-143; 184pp; English.

XX

CC The present sequence is a human NB-ARC and CARD containing protein
CC (NAC) gamma isoform cDNA. NAC gamma isoform represents the NAC splice
CC variant in which both the splice regions are absent in the translated
CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
CC (also referred as NB-ARC domain), a caspase-associated recruitment
CC domain (CARD) and a TIM-barrel-like domain. The caspases, cysteine
CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
CC NAC proteins are used for screening modulators that modulates apoptosis,
CC cytokine production, cytokine receptor signalling and other cellular
CC processes. NAC can act as an immunogen for the production of polyclonal
CC and monoclonal antibodies. It can also be used to diagnose and treat
CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
CC pathologies such as adenocarcinomas and leukaemias.
XX

SQ Sequence 4194 BP; 984 A; 1167 C; 1176 G; 867 T; 0 `other`;

Query Match 74.9%; Score 4078; DB 22; Length 4194;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 4194; Conservative 0; Mismatches 0; Indels 96; Gaps 2;

QY 523 ATGGCTGGCGAGCCTGGGCGCGCTGGCTGTACTTGGAGTTCTCTGAAGAAGAGGAG 582
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 583 CTGAAGGAGTTCAGACTTCTGCCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 CTGAAGGAGTTCAGACTTCTGCCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCGCTCAGCCAGAGAAAGAGTGCGCATGGAGGTGGCTCTGTACCTTGTGGCTCAG 702
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 ACACCGCTCAGCCAGAGAAAGAGTGCGCATGGAGGTGGCTCTGTACCTTGTGGCTCAG 180
QY 703 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTCAGG 762
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTCAGG 240
QY 763 TCACCTGTGCCCCAAAGCCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCTACAGCCCCA 822
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 TCACCTGTGCCCCAAAGCCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCTACAGCCCCA 300
QY 823 AGTGAACCCCACTTGGGGTCTCCAGCAAACCACTCCAACCGAGTGTAAATGCCCTGG 882
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 AGTGAACCCCACTTGGGGTCTCCAGCAAACCACTCCAACCGAGTGTAAATGCCCTGG 360
QY 883 ATCCATGAATTGCGGGGGTCCACCAGGGCTCAGAGAGAGGTTTTTGACACAGCTG 942
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 ATCCATGAATTGCGGGGGTCCACCAGGGCTCAGAGAGAGGTTTTTGACACAGCTG 420
QY 943 CCTGACACATCTGGACGCCCTGGAGAGAATCTCTGCCTCACTCTCTACCAAGCTCTT 1002
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 CCTGACACATCTGGACGCCCTGGAGAGAATCTCTGCCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGTCTCCCGAACCATAGTCTCCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 1062
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 481 CCAAGTCTCCCGAACCATAGTCTCCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGGATCCCACTTAGCCCAAGCTAGCACCCAGAGACAGGAG 1122
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 541 GCAGTGTGGGAGCTGGGGATCCCACTCAGCCCAAGCTAGCACCCAGAGACAGGAG 600
QY 1123 GCTCTGGGACCAATGGCTCTGGATGAACGTCAGGAATTTTACTACACAGAAATCAGA 1182
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 601 GCTCTGGGACCAATGGCTCTGGATGAACGTCAGGAATTTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGAGCGGTGTAGGAACG 1242
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 661 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGAGCGGTGTAGGAACG 720
QY 1243 CCCCCAGGGCACAACAGCTTACGCCCAACCAACCATGGGAGCTTCTGTGAGA 1302
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 721 CCCCCAGGGCACAACAGCTTACGCCCAACCAACCATGGGAGCTTCTGTGAGA 780

Qy	1303	GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362	
Db	781	GAGAGCCTCTGTTCCACATGSCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840	
Qy	1363	CTGCTACTTCTTCAAAAGACCTCACCCACAGAACCAAGATCCCTCGTTCAGAGAGAGCTGG	1422	
Db	841	CTGCTACTTCTTACAAAGACCTCACCCACAGAACCAAGATCCCTCGTTCAGAGAGAGCTGG	900	
Qy	1423	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTCAGATTCAGAGACTTATTTGGCCCA	1482	
Db	901	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTCAGATTCAGAGACTTATTTGGCCCA	960	
Qy	1483	GGCCTGATACCCAAAGAACCTCGCATAGTCACTCGACGGGGGCTGCTGAAATTTGGGAAG	1542	
Db	961	GGCCTGATACCCAAAGAACCTCGCATAGTCACTCGACGGGGGCTGCTGAAATTTGGGAAG	1020	
Qy	1543	TCAACTGCGCCAGCGCAGGTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGACCGC	1602	
Db	1021	TCAACTGCGCCAGCGCAGGTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGACCGC	1080	
Qy	1603	TTCCAGCATGCTTCTACTTTCAGCTGCAGAGAGCTGGCCCCAGTCCAAGTGTGTGAGTCTC	1662	
Db	1081	TTCCAGCATGCTTCTACTTTCAGCTGCAGAGAGCTGGCCCCAGTCCAAGTGTGTGAGTCTC	1140	
Qy	1663	GCTCAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTCTGCT	1722	
Db	1141	GCTCAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTCTGCT	1200	
Qy	1723	AGGCCAGAGCGGCTGCTCTTCATCCTCGATGTTAGATGAGCCAGGATGGGTCTTCAG	1782	
Db	1201	AGGCCAGAGCGGCTGCTCTTCATCCTCGATGTTAGATGAGCCAGGATGGGTCTTCAG	1260	
Qy	1783	GAGCCGAGTTCTGAGCTCTGTCTGCACTGAGCCAGCCACAGCCGGCGGATGCATGCTG	1842	
Db	1261	GAGCCGAGTTCTGAGCTCTGTCTGCACTGAGCCAGCCACAGCCGGCGGATGCATGCTG	1320	
Qy	1843	GGCAGTTTGTGGGGAANAATATACTTCCGAGGCATCCTTCTGTATCAGGCTCGGACC	1902	
Db	1321	GGCAGTTTGTGGGGAANAATATACTTCCGAGGCATCCTTCTGTATCAGGCTCGGACC	1380	
Qy	1903	ACAGCTCTGCAGAACCTTCATTCCTTCTTTGGAGCAGGCAGCTTGGGTAGAGGTCTCTGGG	1962	
Db	1381	ACAGCTCTGCAGAACCTTCATTCCTTCTTTGGAGCAGGCAGCTTGGGTAGAGGTCTCTGGG	1440	
Qy	1963	TTCTCTGAGTCCAGCAGGAAGGAATTTCTTACAGATATTTCACAGATGAAGGCAAGCA	2022	
Db	1441	TTCTCTGAGTCCAGCAGGAAGGAATTTCTTACAGATATTTCACAGATGAAGGCAAGCA	1500	
Qy	2023	ATTAGAGCCTTTAGGTTGGTCAANTCAACAAAGAGCTCTGGGCCCTGCTCTGTGTGCC	2082	
Db	1501	ATTAGAGCCTTTAGGTTGGTCAANTCAACAAAGAGCTCTGGGCCCTGCTCTGTGTGCC	1560	
Qy	2083	TGGGTGCTCGGCTGGCTGCACCTTGCTGATGCAGCAGATGAAGCGGAAGGAAAAAATC	2142	
Db	1561	TGGGTGCTCGGCTGGCTGCACCTTGCTGATGCAGCAGATGAAGCGGAAGGAAAAAATC	1620	
Qy	2143	ACACTGACTTCCAAAGACCCACCAACCTCTGTCTACATTAACCTTGGCCCAAGGCTCTCAA	2202	
Db	1621	ACACTGACTTCCAAAGACCCACCAACCTCTGTCTACATTAACCTTGGCCCAAGGCTCTCAA	1680	
Qy	2203	GCTCAGGCATTTGGAGCCGACGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG	2262	
Db	1681	GCTCAGGCATTTGGAGCCGACGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG	1740	
Qy	2263	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATC	2322	
Db	1741	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATC	1800	
Qy	2323	ATCTCCACCTTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCAATCCCTCTGAGCTACAGC	2382	
Db	1801	ATCTCCACCTTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCAATCCCTCTGAGCTACAGC	1860	

QY	2383	TTCAATTACACCTCTGTGTTTCCAAAGAGTCTTTTGGACAGCAATGTCCTATGCTCTTGGAGGATGAG	2444
DB	1861	TTCAATTACACCTCTGTGTTTCCAAAGAGTCTTTTGGACAGCAATGTCCTATGCTCTTGGAGGATGAG	1920
QY	2443	AAGGGGAGAGGTAAACATTTCTAAATTGCATCATAGATTTTGGAAAGACGCTPAGAAGCATAT	2502
DB	1921	AAGGGGAGAGGTAAACATTTCTAAATTGCATCATAGATTTTGGAAAGACGCTPAGAAGCATAT	1980
QY	2503	GGAAATACATGGCCCTGTTTGGGGCATCAACACACACGTTTTCCTATTGGGCCCTGTTTAAGTGAT	2562
DB	1981	GGAAATACATGGCCCTGTTTGGGGCATCAACACACACGTTTTCCTATTGGGCCCTGTTTAAGTGAT	2040
QY	2563	GAGGGGAGAGAGATPGAGAAACATCTTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTTG	2622
DB	2041	GAGGGGAGAGAGATPGAGAAACATCTTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTTG	2100
QY	2623	ATGCAGTGGGTCGCGTCCCTCGAGCTGCTGCTGCAGCCACACCTCTCTGGAGTCCCTCCAC	2682
DB	2101	ATGCAGTGGGTCGCGTCCCTCGAGCTGCTGCTGCAGCCACACCTCTCTGGAGTCCCTCCAC	2160
QY	2683	TGCTTGTACGAGACTCGGAAACAAAACGTTTCCCTGACACAAAGTGATGGCCCATTTTCGAAGAA	2742
DB	2161	TGCTTGTACGAGACTCGGAAACAAAACGTTTCCCTGACACAAAGTGATGGCCCATTTTCGAAGAA	2220
QY	2743	ATGGGCATGTGTGTAGAAAACAGACATGGAGCTTTTATGTGCACTTTTCTGCATTAATAATTC	2802
DB	2221	ATGGGCATGTGTGTAGAAAACAGACATGGAGCTTTTATGTGCACTTTTCTGCATTAATAATTC	2280
QY	2803	AGCGCCACGCTGAAGAAGCTTTCAGCTGTATTGAGGCAGGCAGCACAGATCAACATGGAGC	2862
DB	2281	AGCGCCACGCTGAAGAAGCTTTCAGCTGTATTGAGGCAGGCAGCACAGATCAACATGGAGC	2340
QY	2863	CCACCATGTGTAGTCTGTTTTCAGTGGGTCCAGTCAACAGATGCCTATTGGCAGATTCTC	2922
DB	2341	CCACCATGTGTAGTCTGTTTTCAGTGGGTCCAGTCAACAGATGCCTATTGGCAGATTCTC	2400
QY	2923	TTCTCCGTCCTCAAGGTCACACAGAAACCTGAAGGAGCTGGACCTAAAGTGGAAACTCGCTG	2982
DB	2401	TTCTCCGTCCTCAAGGTCACACAGAAACCTGAAGGAGCTGGACCTAAAGTGGAAACTCGCTG	2460
QY	2983	AGCCACTCTGCAGTGAAGAGTCTTTGTAGACCCCTGAGACGCCCTCGCTGCCTCTCGGAG	3042
DB	2461	AGCCACTCTGCAGTGAAGAGTCTTTGTAGACCCCTGAGACGCCCTCGCTGCCTCTCGGAG	2520
QY	3043	ACCTCGGCTTGGCTGGCTCTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG	3102
DB	2521	ACCTCGGCTTGGCTGGCTCTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG	2580
QY	3103	CTGAGGCCAACACGACCCCTGACCGAGCTGGAACCTGAGTTTCAATGTGCTACCGGATGCT	3162
DB	2581	CTGAGGCCAACACGACCCCTGACCGAGCTGGAACCTGAGTTTCAATGTGCTACCGGATGCT	2640
QY	3163	GGAGCCAAAACACCTTTGCCAGAGACTGAGACGCGGAGCTGCAAGCTACAGCCACTGCAG	3222
DB	2641	GGAGCCAAAACACCTTTGCCAGAGACTGAGACGCGGAGCTGCAAGCTACAGCCACTGCAG	2700
QY	3223	CTGCTCAGCTGTGSCCTCAGCTCTGACTGCTGCCAGACCTGSCCTCTGTGCTTAGTGCC	3282
DB	2701	CTGCTCAGCTGTGSCCTCAGCTCTGACTGCTGCCAGACCTGSCCTCTGTGCTTAGTGCC	2760
QY	3283	AGCCCCAGCCTGAAGGAGCTAGACCTGACGAGAAACAACTGGATGACGTTGGCGTGCGA	3342
DB	2761	AGCCCCAGCCTGAAGGAGCTAGACCTGACGAGAAACAACTGGATGACGTTGGCGTGCGA	2820
QY	3343	CTGCTCTGTAGGGGCTCAGGCATCTTGCTCTGAAACTCATACGCTGGGGCTGGACCAG	3402
DB	2821	CTGCTCTGTAGGGGCTCAGGCATCTTGCTCTGAAACTCATACGCTGGGGCTGGACCAG	2867
QY	3403	ACAACTCTGAGTGATGATGAGGACGAGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG	3462
DB	2868	-----	2867
QY	3463	CTGCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTAGGGCCTGGATACG	3522

Db 2868 -----GAAACCAAGTGTGATGACCCCTACTGAGGGCTGTGATACG 2907
 QY 3523 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 3582
 Db 2908 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 2967
 QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCTGAGCAAGATCTTCCCAATT 3642
 Db 2968 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCTGAGCAAGATCTTCCCAATT 3027
 QY 3643 GCTGAGATTGACAGGAAGCTCTCCAGAGGTAGTACCGGTGGAATCTTGTGCGTGCCT 3702
 Db 3028 GCTGAGATTGACAGGAAGCTCTCCAGAGGTAGTACCGGTGGAATCTTGTGCGTGCCT 3087
 QY 3703 TCTCTGCTCTCTCAAGGGGACCTGATACGAAGCCCTTTGGGACCTGACGATGACTTCTGG 3762
 Db 3088 TCTCTGCTCTCTCAAGGGGACCTGATACGAAGCCCTTTGGGACCTGACGATGACTTCTGG 3147
 QY 3763 GGGCCACAGGGGCTGTGCTACTAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
 Db 3148 GGGCCACAGGGGCTGTGCTACTAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3207
 QY 3823 CACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
 Db 3208 CACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3267
 QY 3883 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTCTCTGGGTGAGATCAAC 3942
 Db 3268 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTCTCTGGGTGAGATCAAC 3327
 QY 3943 CCACAGCAGAGCTGGATGTGGCAGGGCTCTGCTGACATCAAGCTGAGCTTGAGCT 4002
 Db 3328 CCACAGCAGAGCTGGATGTGGCAGGGCTCTGCTGACATCAAGCTGAGCTTGAGCT 3387
 QY 4003 GTGGAAGCTGTGACCTCCCTCACTTTTGTGGCTCTCCAGGGGGCCATGTGACACATCC 4062
 Db 3388 GTGGAAGCTGTGACCTCCCTCACTTTTGTGGCTCTCCAGGGGGCCATGTGACACATCC 3447
 QY 4063 CTGTTCCAAATGGCCACCTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 4122
 Db 3448 CTGTTCCAAATGGCCACCTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 3507
 QY 4123 GAGCTGCATCAGATGTTCTGGAACCCAGCTCTCTCCCTCTGGGAGTCTCTCTGAAA 4182
 Db 3508 GAGCTGCATCAGATGTTCTGGAACCCAGCTCTCTCCCTCTGGGAGTCTCTCTGAAA 3567
 QY 4183 ATGATCCATAATGCTGCGCTTCAATCCGCTCACTCTGTGGTGTGTTTACCAACCGC 4242
 Db 3568 ATGATCCATAATGCTGCGCTTCAATCCGCTCACTCTGTGGTGTGTTTACCAACCGC 3627
 QY 4243 GTCCATCTTGAGGAGTCACTTCACTCTACCTGATCCCAAGTCACTGCTCCATTCCG 4302
 Db 3628 GTCCATCTTGAGGAGTCACTTCACTCTACCTGATCCCAAGTCACTGCTCCATTCCG- 3686
 QY 4303 AAGGAATCGAGCTCTGCTATCGAACCCCTGGAGAGACAGCTGTTCTCGGAGTTCTAC 4362
 Db 3687 --GGAATCGAGCTCTGCTATCGAACCCCTGGAGAGACAGCTGTTCTCGGAGTTCTAC 3744
 QY 4363 GTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAGACAAGAAAGATGAGACTTG 4422
 Db 3745 GTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAGACAAGAAAGATGAGACTTG 3804
 QY 4423 GTTGGGAGCCCTTGGTGAACAGAGAGATCTCATGCTGCACTACTCTGATCCCTCCA 4482
 Db 3805 GTTGGGAGCCCTTGGTGAACAGAGAGATCTCATGCTGCACTACTCTGATCCCTCCA 3864
 QY 4483 GCCCGATAGCCGTACTCTCACTCTGGATGCCCGCAGTTGCTGACATTTGTGACACAG 4542
 Db 3865 GCCCGATAGCCGTACTCTCACTCTGGATGCCCGCAGTTGCTGACATTTGTGACACAG 3924
 QY 4543 TATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGAGGTGTTCTTTGGACAACTGCAT 4602

Db 3925 TATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGACAAACTGCAT 3984
 QY 4603 GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTTGTGGCTGAGAAACACGAGGCCCCAGC 4662
 Db 3985 GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTTGTGGCTGAGAAACACGAGGCCCCAGC 4044
 QY 4663 CAGATCGGAAGCTGTTGAGCTTGAGCCAGTCTCTGGACCGGAAGTGAAGATGAGACTC 4722
 Db 4045 CAGATCGGAAGCTGTTGAGCTTGAGCCAGTCTCTGGACCGGAAGTGAAGATGAGACTC 4104
 QY 4723 TACCAAGCCCTGAAGGAGACCCATCTCCACCTCATTTATGGAATCTGGGAGAGGGCAGC 4782
 Db 4105 TACCAAGCCCTGAAGGAGACCCATCTCCACCTCATTTATGGAATCTGGGAGAGGGCAGC 4164
 QY 4783 AAAAGGGGACTCTGCCACTCAGCAGCTGA 4812
 Db 4165 AAAAGGGGACTCTGCCACTCAGCAGCTGA 4194

RESULT 8
 AAD02762
 ID AAD02762 standard; cDNA; 4329 BP.
 XX
 AC AAD02762;
 XX
 XX 31-MAY-2001 (first entry)
 XX
 DE Human NB-ARC and CARD containing protein (NAC) delta isoform cDNA.
 XX
 KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 CDS
 Location/Qualifiers
 1..4329
 /tag= a
 /product= "Human NB-ARC and CARD containing protein (NAC)
 delta isoform"
 1..2868
 /tag= b
 /note= "Corresponds to 1-2868 residues of human NAC
 beta isoform (AAD02760)"
 2869..4329
 /tag= c
 /note= "Corresponds to 2962-4422 residues of human NAC
 beta isoform (AAD02760)"
 XX
 WO200116170-A2.
 XX
 XX 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Reed JC;
 XX
 WPI: 2001-183258/18.
 DR P-PSDB; AAY72671.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 XX
 PS Claim 4; Page 148-154; 184pp; English.
 XX
 CC The present sequence is a human NB-ARC and CARD containing protein

CC (NAC) delta isoform cDNA. NAC delta isoform represents the NAC splice
 CC variant in which one of the splice region is absent in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TRIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.
 XX
 SQ Sequence 4329 BP; 1018 A; 1202 C; 1208 G; 901 T; 0 other;

Query Match 72.6%; Score 3952; DB 22; Length 4329;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 4197; Conservative 0; Mismatches 0; Indels 225; Gaps 2;

QY 523 ATGGCTGGCGGAGCTGGGGCGCGCTGCTTACTTGGAGTTCCTGAAGAGGAGGAG 582
 DB 1 ATGGCTGGCGGAGCTGGGGCGCGCTGCTTACTTGGAGTTCCTGAAGAGGAGGAG 60
 QY 583 CTGAAGGAGTTCAGGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
 DB 61 CTGAAGGAGTTCAGGCTTCTGCTCGCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
 QY 643 ACACCGCTCAGCAGAGAAAGAGAGTGGAGTGGAGTGGCTCGTACCTGGTGGCTCAG 702
 DB 121 ACACCGCTCAGCAGAGAAAGAGAGTGGAGTGGAGTGGCTCGTACCTGGTGGCTCAG 180
 QY 703 TATGGGAGCAGGGCGCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 762
 DB 181 TATGGGAGCAGGGCGCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 240
 QY 763 TCACTGTGCGCCCAAGCCAGGAAGGGGAGGCGCACTCTCCCTCACTTCGCCCTACAGCCCA 822
 DB 241 TCACTGTGCGCCCAAGCCAGGAAGGGGAGGCGCACTCTCCCTCACTTCGCCCTACAGCCCA 300
 QY 823 AGTGAACCCCACTGGGGTCTCCAGCAACCCCACTCCAGCGAGTGTAAATGCCCTGG 882
 DB 301 AGTGAACCCCACTGGGGTCTCCAGCAACCCCACTCCAGCGAGTGTAAATGCCCTGG 360
 QY 883 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGGGTTTGGAGACAGCTG 942
 DB 361 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGGGTTTGGAGACAGCTG 420
 QY 943 CTTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT 1002
 DB 421 CTTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT 480
 QY 1003 CCAAGCTCCCAAGACCATGAGTCTCAAGCAGGAGTCAACCAAGCCGCCCAATCCACA 1062
 DB 481 CCAAGCTCCCAAGACCATGAGTCTCAAGCAGGAGTCAACCAAGCCGCCCAATCCACA 540
 QY 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGCCAGAGGAGGAG 1122
 DB 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGCCAGAGGAGGAG 600
 QY 1123 GCTCTGGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTTACTACAGAAATTCAGA 1182
 DB 601 GCTCTGGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTTACTACAGAAATTCAGA 660
 QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCTCCATGGGAGCGGTGTAGGAAAG 1242
 DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCTCCATGGGAGCGGTGTAGGAAAG 720
 QY 1243 CCCCCACGGCCACACCGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA 1302
 DB 721 CCCCCACGGCCACACCGCTTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA 780
 QY 1303 GAGAGCTCTGTTCACATGGCGCTCGAAAAATGAGGATTTTAAACCAAAAAATTCACACAG 1362

DB 781 GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAAAATTCACACAG 840
 QY 1363 CTGCTACTTCTTACAAAGACCTCACCCAGAGAGCCAGAGTCCCTGCTCAAGAGAAGCTGG 1422
 DB 841 CTGCTACTTCTTACAAAGACCTCACCCAGAGAGCCAGAGTCCCTGCTCAAGAGAAGCTGG 900
 QY 1423 CTTGATTAATGTGGAGGAGAAATCGAGGACATTAATTTAGATCAGAGACTTAATTTGGCCCA 1482
 DB 901 CTTGATTAATGTGGAGGAGAAATCGAGGACATTAATTTAGATCAGAGACTTAATTTGGCCCA 960
 QY 1483 GGCCTGGATACCCAAAGACCTCGCATATCATATGCAAGGGGCTGTGGAAATTTGGGAAG 1542
 DB 961 GGCCTGGATACCCAAAGACCTCGCATATCATATGCAAGGGGCTGTGGAAATTTGGGAAG 1020
 QY 1543 TCACACTGGCCAGGCGAGTGAAGAGCCTGGGGAGAGCCAGCTATATGGGACCGC 1602
 DB 1021 TCACACTGGCCAGGCGAGTGAAGAGCCTGGGGAGAGCCAGCTATATGGGACCGC 1080
 QY 1603 TTCCAGCATGTCTTCTACTTTCAGCTGCAAGAGCTGGGCCAGTCCAAGTGGTGAAGTCTC 1662
 DB 1081 TTCCAGCATGTCTTCTACTTTCAGCTGCAAGAGCTGGGCCAGTCCAAGTGGTGAAGTCTC 1140
 QY 1663 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1722
 DB 1141 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1200
 QY 1723 AGCCAGAGCGGCTGCTTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTGAG 1782
 DB 1201 AGCCAGAGCGGCTGCTTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTGAG 1260
 QY 1783 GAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGCGGATGCACTGTCTG 1842
 DB 1261 GAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGCGGATGCACTGTCTG 1320
 QY 1843 GGCAGTTTGTGGGAAAAAATATATCTCCGAGGCACTCTTCTGATCAGCGCTCGGACC 1902
 DB 1321 GGCAGTTTGTGGGAAAAAATATATCTCCGAGGCACTCTTCTGATCAGCGCTCGGACC 1380
 QY 1903 ACAGCTCTGCAAGACCTCATTTCTTTGGAGAGGAGCAGCTTTGGGTAGAGTCTCTGGG 1962
 DB 1381 ACAGCTCTGCAAGACCTCATTTCTTTGGAGAGGAGCAGCTTTGGGTAGAGTCTCTGGG 1440
 QY 1963 TTCTCTGAGTCCAGAGGAGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 2022
 DB 1441 TTCTCTGAGTCCAGAGGAGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 1500
 QY 2023 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
 DB 1501 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560
 QY 2083 TGGGTGCTGGCTGGCTGGCTGCACTTGCCTGATGAGCAGAGATGAAGCGGAAGAAACATC 2142
 DB 1561 TGGGTGCTGGCTGGCTGGCTGCACTTGCCTGATGAGCAGAGATGAAGCGGAAGAAACATC 1620
 QY 2143 ACACCTGACTTCCAAAGACCCACCAACCTCTGTCTTACATTTACTTTGCCAGGCTCTCCAA 2202
 DB 1621 ACACCTGACTTCCAAAGACCCACCAACCTCTGTCTTACATTTACTTTGCCAGGCTCTCCAA 1680
 QY 2203 GCTAGCATTGGGACCCAGCTCAGAGACTCTGCTCTCTGGCTGTGGGAGGATCTGG 2262
 DB 1681 GCTAGCATTGGGACCCAGCTCAGAGACTCTGCTCTCTGGCTGTGGGAGGATCTGG 1740
 QY 2263 CAAAAAAGACCTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGGGCGCATC 2322
 DB 1741 CAAAAAAGACCTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGGGCGCATC 1800
 QY 2323 ATCTCCACTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
 DB 1801 ATCTCCACTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860
 QY 2383 TTCATTCACTCTGTTTCAAGAGTTCCTTTCAGCAATGTCTTATGTCTTGGAGGATGAG 2442
 DB 1861 TTCATTCACTCTGTTTTCAGAGTTCCTTTCAGCAATGTCTTATGTCTTGGAGGATGAG 1920

Db 3988 CTGATCCCTCCAGCCCGCATAGCCGTACCTTACCTCTGGATGCCCGCAGTTGCTGCAC 4047
 QY 4531 TTGTGGACCATGATGAGGAGGAGTATAGCCGAGTGACATCGGTGAGGTGCTTG 4590
 Db 4048 TTGTGGACCATGATGAGGAGGAGTATAGCCGAGTGACATCGGTGAGGTGCTTG 4107
 QY 4591 GACAAATGTCATGACGAGGTGCTGAGCCAGGAGCAGTACGAGGGGTGCTGCTGAGAAC 4650
 Db 4108 GACAAATGTCATGACGAGGTGCTGAGCCAGGAGCAGTACGAGGGGTGCTGCTGAGAAC 4167
 QY 4651 ACAGGCCAGCAGATGCGGAAGCTGTTACGTTGAGTCAGTCCTGAGCCGGAAGTGC 4710
 Db 4168 ACAGGCCAGCAGATGCGGAAGCTGTTACGTTGAGTCAGTCCTGAGCCGGAAGTGC 4227
 QY 4711 AAAGTGGACTTACCAAGCCCTGAGGAGACCCATCTCCTCACTATTATGAACTCTGG 4770
 Db 4228 AAAGTGGACTTACCAAGCCCTGAGGAGACCCATCTCCTCACTATTATGAACTCTGG 4287
 QY 4771 GAGAGGGCAGCAAAAGGACTCTCTGCCACTCAGCAGCTGA 4812
 Db 4288 GAGAGGGCAGCAAAAGGACTCTCTGCCACTCAGCAGCTGA 4329

RESULT 9
 AAD02764
 ID AAD02764 standard; cDNA; 4556 BP.
 XX AAD02764;
 XX 31-MAY-2001 (first entry)
 XX Human NAC beta isoform-CARD-X1 chimeric cDNA.
 DE Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
 KW chimeric protein; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 CDS 1..4365
 FT /tag= a
 FT /product= "Human NAC beta isoform-CARD-X1 chimera"
 FT misc_feature 1..3234
 FT /tag= b
 FT /note= "Corresponds to 1-3234 residues of human
 NAC beta isoform (AAD02760)"
 FT misc_feature 3235..4362
 FT /tag= c
 FT /note= "Corresponds to 166-1293 residues of human
 CARD-X (AAD02763)"
 XX
 PN WO200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 XX (BURN-) BURNHAM INST.
 PA
 XX
 PI Reed JC;
 XX
 XX WPI: 2001-183258/18.
 DR P-PSDB; AAY72673.
 XX
 XX Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of

PT apoptosis -
 XX Disclosure; Page 162-168; 184pp; English.
 XX The present sequence is a human NB-ARC and CARD containing protein (NAC)
 CC beta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
 CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
 CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
 CC domain. CARD-X protein comprises a caspase-associated recruitment domain
 CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
 CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
 CC for screening modulators that modulates apoptosis, cytokine production,
 CC cytokine receptor signalling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory
 CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukaemias.
 XX
 SQ Sequence 4556 BP; 1100 A; 1223 C; 1256 G; 977 T; 0 other;
 Query Match 62.5%; Score 3400.4; DB 22; Length 4556;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 3555; Conservative 0; Mismatches 236; Indels 3; Gaps 1;
 QY 523 ATGGCTGGCGGAGCTGGGGCGGCTGGCTGTACTTTGGAGTTCTCTGAAGAAGAGGAG 582
 Db 1 ATGGCTGGCGGAGCTGGGGCGGCTGGCTGTACTTTGGAGTTCTCTGAAGAAGAGGAG 60
 QY 583 CTGAAGAGATTCCAGCTTCTGCTGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
 Db 61 CTGAAGAGATTCCAGCTTCTGCTGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
 QY 643 ACACCCGCTCAGCAGAGAGAGAGAGTGGCTGGAGTGGCTCTACTTGGTGTCTCAG 702
 Db 121 ACACCCGCTCAGCAGAGAGAGAGAGTGGCTGGAGTGGCTCTACTTGGTGTCTCAG 180
 QY 703 TATGGGAGCAGCGGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGTGAAG 762
 Db 181 TATGGGAGCAGCGGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGTGAAG 240
 QY 763 TCAGTGTGGGCCAAGCCCAAGAGAGGGGCGAGCCACTCTCCCTCATTTCCCTTACAGCCCA 822
 Db 241 TCAGTGTGGGCCAAGCCCAAGAGAGGGGCGAGCCACTCTCCCTCATTTCCCTTACAGCCCA 300
 QY 823 AGTGAACCCCACTGGGGTCTCCAGCCCAACCCACTCCACGCGAGTGTAAATGCCCTGG 882
 Db 301 AGTGAACCCCACTGGGGTCTCCAGCCCAACCCACTCCACGCGAGTGTAAATGCCCTGG 360
 QY 883 ATCCATGAATTGCCGGGGGTGCAACCAGGGCTCAGAGAGAAGGGTTTGGAGACAGCTG 942
 Db 361 ATCCATGAATTGCCGGGGGTGCAACCAGGGCTCAGAGAGAAGGGTTTGGAGACAGCTG 420
 QY 943 CTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCTCATCTCTCTACCAAGCTCTT 1002
 Db 421 CTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCTCATCTCTCTCTACCAAGCTCTT 480
 QY 1003 CCAAGCTCCCGAGCAGTGTCTCCAGCCAGGAGTCAACCAAGCCCGCCACATCCACA 1062
 Db 481 CCAAGCTCCCGAGCAGTGTCTCCAGCCAGGAGTCAACCAAGCCCGCCACATCCACA 540
 QY 1063 GCAGTGTGGGAGCTGGGGATCCCCACTCAGCCCGCTAGCAACCAGAGAGAGGAG 1122
 Db 541 GCAGTGTGGGAGCTGGGGATCCCCACTCAGCCCGCTAGCAACCAGAGAGAGGAG 600
 QY 1123 GCTCTGGGACCCCAATGGCTCTGGATGAAGCTCAGGAATTTACTACAGAAATCAGA 1182
 Db 601 GCTCTGGGACCCCAATGGCTCTGGATGAAGCTCAGGAATTTACTACAGAAATCAGA 660
 QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCGGGCCCATCGGAGCGGTGTAGGAACG 1242
 Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGCGGGCCCATCGGAGCGGTGTAGGAACG 720
 QY 1243 CCCCCACAGGGGCACACAGGCTTACAGCCCGCCACACCCATGGGAGCTTCTGTGAGA 1302

[illegible]

Db	1801	ATCTCCACCTTCTTTGAAGATGGGTATTTCTTCAAGAGCACCCCAATCCCTCTGAGCTACAGC	1861
Qy	2383	TTCAATTCACCTCTGTCTTTCCAAAGAGTTCTTTTGCAGCAATGTCTTATGTCTTGGAGGATGAG	2442
Db	1861	TTCAATTCACCTCTGTCTTTCCAAAGAGTTCTTTTGCAGCAATGTCTTATGTCTTGGAGGATGAG	1920
Qy	2443	AAGGGGAGAGGTAAACATCTTAATTGCATCATATAGATTTTGAAAAAGCGCTTAGAAGCATAT	2502
Db	1921	AAGGGGAGAGGTAAACATCTTAATTGCATCATATAGATTTTGAAAAAGCGCTTAGAAGCATAT	1980
Qy	2503	GGAAATACATGGCCTGTTTGGGGCATCAACACACACGTTTCTTATTTGGSCCTGTTTAAAGTGAT	2562
Db	1981	GGAAATACATGGCCTGTTTGGGGCATCAACACACACGTTTCTTATTTGGSCCTGTTTAAAGTGAT	2040
Qy	2563	GAGGGGAGAGAGATGAGAAACATCTTTTCACTGCGCGTGTCTCAAGGGAGGAAACCTG	2622
Db	2041	GAGGGGAGAGAGATGAGAAACATCTTTTCACTGCGCGTGTCTCAAGGGAGGAAACCTG	2100
Qy	2623	ATGCAGTGGGTCCCGTCCCTGACGTGCTGCTGCAGGCCACACCTCTCTGGAGTCCCTCCAC	2682
Db	2101	ATGCAGTGGGTCCCGTCCCTGACGTGCTGCTGCAGGCCACACCTCTCTGGAGTCCCTCCAC	2160
Qy	2683	TGCTTGTACGAGACTCCGGAACAAAACGTTCTCTGCACAAAGTGATGGGCCCATTTTCAAGAA	2742
Db	2161	TGCTTGTACGAGACTCCGGAACAAAACGTTCTCTGCACAAAGTGATGGGCCCATTTTCAAGAA	2220
Qy	2743	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATTTTCTGCATTTAAATTC	2802
Db	2221	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATTTTCTGCATTTAAATTC	2280
Qy	2803	AGCCGCCACGTGAAGAAAGCTTTCAGCTGATTTAGGGGACGGCAGCACAGATCAACATGGAGC	2862
Db	2281	AGCCGCCACGTGAAGAAAGCTTTCAGCTGATTTAGGGGACGGCAGCACAGATCAACATGGAGC	2340
Qy	2863	CCACCAATGGTAGTCTCTGTTTCAGTGGGTCCCAGTCCAGATCCATTCGTCAGATTTCTC	2922
Db	2341	CCACCAATGGTAGTCTCTGTTTCAGTGGGTCCCAGTCCAGATCCATTCGTCAGATTTCTC	2400
Qy	2923	TTCTCCGTCCTCAAGGTCACACAGAAACCTGAAAGAGCTGGACCTTAACTGGAAACCTCGCTG	2982
Db	2401	TTCTCCGTCCTCAAGGTCACACAGAAACCTGAAAGAGCTGGACCTTAACTGGAAACCTCGCTG	2460
Qy	2983	AGCCACTCTGACGTGAAGAGTCTTTTGTAAAGACCTCTGAGACCGCCCTCGCTCGCTCGGAG	3042
Db	2461	AGCCACTCTGACGTGAAGAGTCTTTTGTAAAGACCTCTGAGACCGCCCTCGCTCGCTCGGAG	2520
Qy	3043	ACCTCGGGTTGGCTGCTGTGGCTCACAGCTGAGGACCTGCAAGGACCTTGCCCTTTGGG	3102
Db	2521	ACCTCGGGTTGGCTGCTGTGGCTCACAGCTGAGGACCTGCAAGGACCTTGCCCTTTGGG	2580
Qy	3103	CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT	3162
Db	2581	CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT	2640
Qy	3163	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG	3222
Db	2641	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG	2700
Qy	3223	CTGCTCAGCTGTGGCCTTCAGTCTGACTGCTGCCAGACCTTGGCCTCTGTGCTTAGTGCC	3282
Db	2701	CTGCTCAGCTGTGGCCTTCAGTCTGACTGCTGCCAGACCTTGGCCTCTGTGCTTAGTGCC	2760
Qy	3283	AGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAAACCTGGATGACGTTGGCGTGCGA	3342
Db	2761	AGCCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAAACCTGGATGACGTTGGCGTGCGA	2820
Qy	3343	CTGCTCTGTGAGGGGCTCAGGCATCTCTGCCTGCAAACTCATACGCTGGGCTGGACCG	3402
Db	2821	CTGCTCTGTGAGGGGCTCAGGCATCTCTGCCTGCAAACTCATACGCTGGGCTGGACCG	2880
Qy	3403	ACAACCTGTAGTGTAGATGAGCGAGAACTGAGGGCCCTTGGAGCGAGGAAACCTTCAG	3462
Db	2881	ACAACCTGTAGTGTAGATGAGCGAGAACTGAGGGCCCTTGGAGCGAGGAAACCTTCAG	2940

QY 3463 CTGCTCATCTTTCAGCAGACGAAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATAG 3522
 DB |||||
 QY 2941 CTGCTCATCTTTCAGCAGACGAAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATAG 3000
 DB |||||
 QY 3523 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
 DB |||||
 QY 3001 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3060
 DB |||||
 QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGAGCTGAGCAAGATCTTCCCAATT 3642
 DB |||||
 QY 3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGAGCTGAGCAAGATCTTCCCAATT 3120
 DB |||||
 QY 3643 GCTGAGATTTGCAGAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTTGTGCGTGCCT 3702
 DB |||||
 QY 3121 GCTGAGATTTGCAGAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTTGTGCGTGCCT 3180
 DB |||||
 QY 3703 TCTCTGCTCTCAAGGGACCTGTCATCAAGAGCCCTTTGGGACTGACGATGACTTCTGG 3762
 DB |||||
 QY 3181 TCTCTGCTCTCAAGGGACCTGTCATCAAGAGCCCTTTGGGACTGACGATGACTTCTGG 3240
 DB |||||
 QY 3763 GGGCCACCGGGCTGTGCTACTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 3822
 DB |||||
 QY 3241 GGGCCCTGAAGGAAATGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGAT 3300
 DB |||||
 QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACCGGCTCTGTCTTGTGATGAGA 3882
 DB |||||
 QY 3301 TGGTTCCTCCACTGCTGCTGATCTGTGTTGCTGATGATGATGATGATGATGATGATGAT 3360
 DB |||||
 QY 3883 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTTCCTGG---GTGATGTC 3939
 DB |||||
 QY 3361 GATGAGTTCACAGTGCAGTTCGCTTGTGTTGCTTCTGAGTTCAGCCTGGACCTG 3420
 DB |||||
 QY 3940 AACCCACAGCAGCTGGATGTTGGGAGGCGCTCTGCTGAGCATCAAGCTGAGCTGGA 3999
 DB |||||
 QY 3421 CAGCACCATGAACAGTGGCTGTGGGCGGCCCTTGTGTTGATGTCATGTCAGAGCCAGAG 3480
 DB |||||
 QY 4000 GCTGTGGAAGCTGTGCACCTCCCTCACTTGTGTGGCTCTCCAAAGGGGCGCATGTGGACACA 4059
 DB |||||
 QY 3481 GAGGCTGTGCGGAATCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3540
 DB |||||
 QY 4060 TCCCTGTTCCAAATGCGCCACTTTAAAGAGGAGGGATGCTCTGAGAGGAGGAGGAGGAGG 4119
 DB |||||
 QY 3541 TCTGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3600
 DB |||||
 QY 4120 GTGGAGCTGCATCAGATGTTCTGGAACCCAGCTTCTCCCTTGGGACTTCTCTG 4179
 DB |||||
 QY 3601 GTGGAGCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3660
 DB |||||
 QY 4180 AAAATGATCCATAATGCGCTGCTTCACTTCCGCTCACCTCTGCTGCTGCTGCTGCTGCTG 4239
 DB |||||
 QY 3661 CGGATCGGAGTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720
 DB |||||
 QY 4240 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4299
 DB |||||
 QY 3721 CACCCCGCCCGAAGATTAAGTTCCACTTGTACTTGTCTGCTGCTGCTGCTGCTGCTGCTG 3780
 DB |||||
 QY 4300 CGAAGGAAGTGA 4313
 DB |||||
 QY 3781 ACAAGGCGGATAGA 3794

RESULT 10
 AAD02765
 ID AAD02765 standard; cDNA; 4466 BP.

XX
 AC AAD02765;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Human NAC gamma or delta isoform-CARD-X1 chimeric cDNA.
 XX
 KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;

KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
 KW chimeric protein; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 CDS 1..4275
 FT /tag= a
 FT /product= "Human NAC gamma/delta isoform-CARD-X1 chimera"
 FT 1..3144
 FT /tag= b
 FT /note= "Corresponds to 1-3144 residues of human
 NAC gamma (AAD02761) or delta (AAD02762) isoform"
 FT 3145..4272
 FT /tag= c
 FT /note= "Corresponds to 166-1293 residues of human
 CARD-X (AAD02763)"
 XX WO200116170-A2.
 XX 08-MAR-2001.
 XX 01-SEP-2000; 2000WO-US24152.
 XX 01-SEP-1999; 99US-0388221.
 XX (BURN-) BURNHAM INST.
 XX Reed JC;
 XX WPI: 2001-183258/18.
 XX P-PSDB; AAY72674.
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 XX Disclosure; Page 173-178; 184pp; English.
 PS The present sequence is a human NB-ARC and CARD containing protein (NAC)
 XX gamma or delta isoform-CARD-X1 chimeric cDNA. NAC protein comprises a
 CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
 CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
 CC domain. CARD-X protein comprises a caspase-associated recruitment domain
 CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
 CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
 CC for screening modulators that modulates apoptosis, cytokine production,
 CC cytokine receptor signalling and other cellular processes. They can act
 CC as an immunogen for the production of polyclonal and monoclonal
 CC antibodies. They can also be used to diagnose and treat inflammatory
 CC disorders such as sepsis, fibrosis and arthritis and cancer pathologies
 CC such as adenocarcinomas and leukaemias.
 XX Sequence 4466 BP; 1075 A; 1201 C; 1227 G; 963 T; 0 other;
 SQ
 Query Match 59.0%; Score 3210.4; DB 22; Length 4466;
 Best Local Similarity 91.3%; Pred. No. 0;
 Matches 3465; Conservative 0; Mismatches 236; Indels 93; Gaps 2;
 QY 523 ATGGCTGGCGGAGCTGGGCGGCTGGCTGCTTACTTGGAGTTCCTCAAGAGGAGGAG 582
 DB |||||
 DB 1 ATGGCTGGCGGAGCTGGGCGGCTGGCTGCTTACTTGGAGTTCCTCAAGAGGAGGAG 60
 QY 583 CTGAGGAGTTCAGACTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
 DB |||||
 DB 61 CTGAGGAGTTCAGACTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
 QY 643 ACACCGCTCAGCCAGAGAAACAGGTGGATGAGGTGGCTCTGTACTGCTGCTCTAG 702
 DB |||||
 DB 121 ACACCGCTCAGCCAGAGAGACAGTGGATGGAGTGGCTCTGTACTGCTGCTCTAG 180

QY	703	TATGGGGAGCAGCGGGCTTGGGACCTAGCCCTCCATATACCTGGAGCAGATGGGGTGGG	762	1783	GAGCCGAGTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG	1842
Db	181	TATGGGGAGCAGCGGGCTTGGGACCTAGCCCTCCATATACCTGGAGCAGATGGGGTGGG	240	1261	GAGCCGAGTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG	1320
QY	763	TCACTGTGCGCCCAAGCCCAAGAGGGGAGGGCCACTCTCCCTCATTCCTCCCTACAGCCCA	822	1843	GGCAGTTTGTGGGAAAAAATACTTCCCGAGGCATCTCTCTCATCACGGCTCGGACC	1902
Db	241	TCACTGTGCGCCCAAGCCCAAGAGGGGAGGGCCACTCTCCCTCATTCCTCCCTACAGCCCA	300	1321	GGCAGTTTGTGGGAAAAAATACTTCCCGAGGCATCTCTCTCATCACGGCTCGGACC	1380
QY	823	AGTGAACCCACCTCGGGGTCTCCAGCCAAACCCACCTCCACCGCAGTGTAAATGCCCTGG	882	1903	ACAGCTCTGCAGAACCTCAITCTCTTTTGGAGCAGGCACGTGGGTAGAGTCTCTGGGG	1962
Db	301	AGTGAACCCACCTCGGGGTCTCCAGCCAAACCCACCTCCACCGCAGTGTAAATGCCCTGG	360	1381	ACAGCTCTGCAGAACCTCAITCTCTTTTGGAGCAGGCACGTGGGTAGAGTCTCTGGGG	1440
QY	883	ATCCATGAATTCGCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGAGACAGCTG	942	1963	TTCTCTGAGTCCAGAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	2022
Db	361	ATCCATGAATTCGCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGAGACAGCTG	420	1441	TTCTCTGAGTCCAGAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	1500
QY	943	CCTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTACCAAGCTCTT	1002	2023	ATTAGAGCCTTTAGGTGTGTCAAATCAAAACAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
Db	421	CCTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTACCAAGCTCTT	480	1501	ATTAGAGCCTTTAGGTGTGTCAAATCAAAACAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
QY	1003	CCAAAGTCTCCAGACCATGAGTCTCCAAAGCAGGAGTCAACCAAGCCGCCACATCCACA	1062	2083	TGGGTGTCTGGCTGGCCCTTGCACTTGGCCTGATGACAGCAGATGAAGCGGAAGAAAACTC	2142
Db	481	CCAAAGTCTCCAGACCATGAGTCTCCAAAGCAGGAGTCAACCAAGCCGCCACATCCACA	540	1561	TGGGTGTCTGGCTGGCCCTTGCACTTGGCCTGATGACAGCAGATGAAGCGGAAGAAAACTC	1620
QY	1063	GCAGTCTGGGAGCTGGGGATCCCAACCTCAGCCAGCCTTAGCACCCAGAGAGCAGGAG	1122	2143	ACACTGACTTCCAAAGCACCACCAACCCCTCTGTCTACATTTACCTTTGCCAGGCTCTCAA	1680
Db	541	GCAGTCTGGGAGCTGGGGATCCCAACCTCAGCCAGCCTTAGCACCCAGAGAGCAGGAG	600	2203	GCTCAGACCATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGGGCATCTGG	2262
QY	1123	GCTCTGGAGCCCAATGGGCTCTGGATGAACGTCAGGAAATTTACTACAGAAATCAGA	1182	1681	GCTCAGACCATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGGGCATCTGG	1740
Db	601	GCTCTGGAGCCCAATGGGCTCTGGATGAACGTCAGGAAATTTACTACAGAAATCAGA	660	2263	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	2322
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCAATGGCGGTGGTAGGAAGC	1242	1741	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC	1800
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCAATGGCGGTGGTAGGAAGC	720	2323	ATCTCCACTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTATCCCTCTGAGCTACAGC	2382
QY	1243	CCCCACAGCGCACACAGCCTTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA	1302	1801	ATCTCCACTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTATCCCTCTGAGCTACAGC	1860
Db	721	CCCCACAGCGCACACAGCCTTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA	780	2383	TTCAATTCACCTGTGTTTCCAGAGTCTTTTTCAGCAATGTCTCTGTGAGGATGAG	2442
QY	1303	GAGAGCCTCTGTTCACATGGCCCTCGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362	1861	TTCAATTCACCTGTGTTTCCAGAGTCTTTTTCAGCAATGTCTCTGTGAGGATGAG	1920
Db	781	GAGAGCCTCTGTTCACATGGCCCTCGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840	2443	AAGGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTGGAAAAAGACGCTAGAAGCATAT	2502
QY	1363	CTGCTACTTCTACAGAGCCTCACCCAGAGCCAGAGTCCCTGGTCAAGAGAGCTGG	1422	1921	AAGGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTGGAAAAAGACGCTAGAAGCATAT	1980
Db	841	CTGCTACTTCTACAAAGACCTCACCCAGAAGCCAGATCCCCCTGGTCAAGAGAAGCTGG	900	2503	GGAAATACATGGCCTGTTTGGGGCATCAACCAACACGTTTCTATTTGGGCTGTAAAGTGAT	2562
QY	1423	CCTGATTATGTGGAGAGAAATCGAGGACATTTTAAATTTGAGATCAGAGACTTATTTGGCCCA	1482	1981	GGAAATACATGGCCTGTTTGGGGCATCAACCAACACGTTTCTATTTGGGCTGTAAAGTGAT	2040
Db	901	CCTGATTATGTGGAGAGAAATCGAGGACATTTTAAATTTGAGATCAGAGACTTATTTGGCCCA	960	2563	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGCTGTCTCAGGGGAGGACCTG	2622
QY	1483	GGCTGGATACCAAGAACTCGCATAGTCAATCTGAGGGGCTGCTGGAATTTGGGAAG	1542	2041	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGCTGTCTCAGGGGAGGACCTG	2100
Db	961	GGCTGGATACCAAGAACTCGCATAGTCAATCTGAGGGGCTGCTGGAATTTGGGAAG	1020	2623	ATGCAAGTGGTCCCGTCCCTGCACTGTCTGTCAGCCACACTCTCTGAGTCCCTCCAC	2682
QY	1543	TCAACACTGGCCAGGAGTGAAGGAGCCTGGGGGAGGCGCAGCTGTATGGGACCGC	1602	2101	ATGCAAGTGGTCCCGTCCCTGCACTGTCTGTCAGCCACACTCTCTGAGTCCCTCCAC	2160
Db	1021	TCAACACTGGCCAGGAGTGAAGGAGCCTGGGGGAGGCGCAGCTGTATGGGACCGC	1080	2683	TGCTTTGACGAGACTCGGAACAAAAAGTTCCTGTGACAAAGTATGAGTCCCAATTTGGAAGAA	2742
QY	1603	TTCCAGATGTCTTCTACTTTCAGCTGCAGAGCTGGCCCGCAGTCCCAAGGTGGTGTCTC	1662	2161	TGCTTTGACGAGACTCGGAACAAAAAGTTCCTGTGACAAAGTATGAGTCCCAATTTGGAAGAA	2220
Db	1081	TTCCAGATGTCTTCTACTTTCAGCTGCAGAGCTGGCCCGCAGTCCCAAGGTGGTGTCTC	1140	2743	ATGGGATGTGTAGAAAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC	2802
QY	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGCGTCCCAATTAGACAGATCTGTCT	1722	2221	ATGGGATGTGTGTAGAAAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC	2280
Db	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGCGTCCCAATTAGACAGATCTGTCT	1200	2803	AGCCGCGCAGTGAAGAGCTTTCAGCTGATTTAGGGCAGGCGACACAGATCAACATGGAGC	2862
QY	1723	AGGCCAGAGCGGCTCTTTCATCTTCGATGGTGTAGATGAGCCAGATGGGTCTTGCAG	1782	2281	AGCCGCGCAGTGAAGAGCTTTCAGCTGATTTAGGGCAGGCGACACAGATCAACATGGAGC	2340
Db	1201	AGGCCAGAGCGGCTCTTTCATCTTCGATGGTGTAGATGAGCCAGATGGGTCTTGCAG	1260	2863	CCCAACCATGGTGTCTCTTTCAGTGGGTCCCAAGTCAAGATGCTCTATTGGCAGATTCTC	2922

[illegible]

D	b		3331	CAGCACCATGAAACAGTGGCTGTGGGGCGGCCCTTTGTTGATGTCACTGCAGAGCCAGAG	3399
Q	y		4000	GCTGTGAAGCTGTGCACCTCCCTCACTTTTGTGGCTCTCCAAAGGGGGCCCATGTGGACACA	4059
D	b		3391	GAGGCTGTGCGCGAATCCACCTCCCACACTTCATCTCCCTCCAAGGTGAGGTGGACGTC	3450
Q	y		4060	TCCCTGTTTCAAATGGGCCACTTTTAAGAGGAGGGGATGCTCTTGAGAGAAGCCAGCCAGG	4119
D	b		3451	TCCTGTGTTTCTCGTTGCCCATTTTAAGAAATGAAGGGATGGTCTCTGGAGCATCCAGGCCGG	3510
Q	y		4120	GTGGAGCTGCATCACATAGTTCCTTGGAAAAACCCAGAGCTTCCTCCCTTGGGAGTCCCTCCTG	4179
D	b		3511	GTGGAGCCTTTCTATGCTGTCTCTGGAAAGCCCCAGACTTCTCTCTGATGGGCATCCTTGCTG	3570
Q	y		4180	AAAATGATCCATAATGCGCTTCGCGCTTCATTTCCTCCGTCACCTCTCTGTGGTGTGCTTTTACCAC	4239
D	b		3571	CGGATCCAGTGGGACTCGCCTCTCCATCCCATCACCTTCCAACACATTGATCTATTAT	3630
Q	y		4240	CGGCTCCATCCTGAGGAAGTCACCTTCCACCTCTACTCTGATGCCAAGTGACTGCTCCATT	4299
D	b		3631	CACCCCCACCCGAAAGATATTAAAGTTCACACTTGTACCTTGCCCCAGCAGCGCTTGCTA	3690
Q	y		4300	CGGNAGGAAGTGG A 4313 	
D	b		3691	ACAAAGGCGATAGA 3704 	
RESULT 11					
AAA78392					
I	D		AAA78392	standard; cDNA; 2657 BP.	
X	X		AC	AAA78392;	
X	X		AC		
X	X		DT	20-NOV-2000 (first entry)	
X	X		XX	Human secreted protein gene 12 SEQ ID NO:22.	
X	X		DE		
K	W		KW	Human; secreted protein; cytostatic; antianaemic; antidiabetic; antiinflammatory; ophthalmological; antirheumatic; antiarthritic; antiporiatic; antiangiogenic; cardiant; anti-HIV; nootropic; neuroprotective; antimicrobial; antiparkinsonian; cancer; immune system disorder; angiogenesis; hyperproliferative disorder; cardiovascular disorder; apoptosis; neurological disease; infectious disease; wound healing; ss.	
X	X		KW	Homo sapiens.	
X	X		OS		
X	X		FN	WO2000035937-Al.	
X	X		PD	22-JUN-2000.	
X	X		PF	16-DEC-1999; 99WO-US299950.	
X	X		PR	17-DEC-1998; 98US-0112809.	
X	X		PR	18-DEC-1998; 98US-0113006.	
X	X		XX	(HUMA-) HUMAN GENOME SCI INC.	
X	X		PA		
X	X		FI	Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;	
X	X		PI	Duan DR, Moore PA, Shi Y, Lafleur DW, Oleen HS, Florence K;	
X	X		DR	WPI; 2000-431566/37. P-PSDB; AAB24448.	
X	X		XX		
P	T		PT	Forty seven human nucleic acids encoding secreted proteins, useful in the treatment, prevention and diagnosis of cancers, disorders of the immune system, angiogenesis disorders, neurological diseases and hyperproliferative disorders - Claim 1; Page 449-450; 562pp; English.	
X	X		XX		
C	C		CC	The polynucleotide sequence given in AAA78381 to AAA78432 encode the human secreted proteins given in AAB24437 to AAB24604. Human secreted	

Db 1867 GCTGTTTCAGCTTGAGCCAGTCTCTGGACCGGAAGTGCAAGATGGACTCTCAAGCCCT 1926
Qy 4734 GAAGGAGACCCATCTCCTACCTCATATGAACTCTGGGAGAGGGCAGCAAAAGGACT 4793
Db 1927 GAAGGAGACCCATCTCTCA-CTCATTTAGAACTCTGGGAGAGGGCAGCAAAAGGACT 1985
Qy 4794 CTGCGCACTCAGCAGCTGAAGTATCAACACCGCCCTTCAGCCCTTCAGTCTGGCTTTGG 4853
Db 1986 CTGCGCACTCAGCAGCTGAAGTATCAACACTAGCCCTTCAGCCCTTCAGTCTGGCTTTGG 2045
Qy 4854 CTGACCCCTCTTTGGTCTCAGTTCTCTCTGCAAAAGTGGCCATCTCGTTTGCCT 4913
Db 2046 CTGACCCCTCTTTGGTCTCAGTTCTCTCTGCAAAAGTGGCCATCTCGTTTGCCT 2105
Qy 4914 TCAGCAGCTAAAGTAATGAACTTTGATGATGCTTTGCTGGGCATTTATGTTCATGCC 4973
Db 2106 TCAGCAGCTAAAGTAATGAACTTTGATGATGCTTTGCTGGGCATTTATGTTCATGCC 2165
Qy 4974 AGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATC 5033
Db 2166 AGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGAATGTCCATC 2225
Qy 5034 TGGAGCTGGCAAGACCCCTCGACACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCAA 5093
Db 2226 TGGAGCTGGCAAGACCCCTCGACACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCA 2285
Qy 5094 GCTAGAGCCCTCCGATCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACAT 5153
Db 2286 AGCCTAGAGCCTCCGATCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACAT 2345
Qy 5154 TTGCTCTCGCTGTGTACAGGTGAGCCCAAAATTTGGGTTTTCAGCGTGGGAGGCCAG 5213
Db 2346 TTGCTCTCGCTGTGTACAGGTGAGCCCAAAATTTGGGTTTTCAGCGTGGGAGGCCAG 2405
Qy 5214 TGGATTCTTTGGCTTTGTACAGGAAGATCTACAAGAGCAAGCCCAACAGAGTAAGTGGAG 5273
Db 2406 TGGATTCTTTGGCTTTGTACAGGAAGATCTACAAGAGCAAGCCCAACAGAGTAAGTGGAG 2465
Qy 5274 GAAGTTTATTCAGAAATAAAGAGTATCACAGCTCTTTTAGAATTTGTAGCAGGCTT 5333
Db 2466 GAAGTTTATTCAGAAATAAAGAGTATCACAGCTCTTTTAGAATTTGTAGCAGGCTT 2525
Qy 5334 TCCAGTTTATTCAGAAACCCCTATAAATTTAAATTTTACTTTAATTTTAAAGATT 5393
Db 2526 TCCAGTTTATTCAGAAACCCCTATAAATTTAAATTTTACTTTAATTTTAAAGATT 2585
Qy 5394 AAAAAATCAAAAAAGAAAAATGAAAAATAGAAATAGAAATAGAAATAGAAATAGAAAT 5444
Db 2586 AAAAAATCAAAAAAGAAAAATGAAAAATAGAAATAGAAATAGAAATAGAAATAGAAAT 2636

RESULT 12

AAL01744

ID AAL01744 standard; cdna; 578 BP.

XX AAL01744;

XX 21-NOV-2001 (first entry)

DE Human reproductive system related antigen cdna SEQ ID NO: 1745.

XX Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
OS Homo sapiens.
XX WO20015320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01339.

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.

PR	02-OCT-2000;	2000US-0237037.	XX	Claim 1; SEQ ID NO 1745; 1297bp + Sequence Listing; English.
PR	02-OCT-2000;	2000US-0237038.	PS	
PR	02-OCT-2000;	2000US-0237039.	XX	
PR	02-OCT-2000;	2000US-0237040.	CC	The present invention provides the protein and coding sequences of a
PR	13-OCT-2000;	2000US-0239335.	CC	number of human reproductive system related antigens. These can be used
PR	13-OCT-2000;	2000US-0239337.	CC	in the prevention and treatment of reproductive system disorders,
PR	20-OCT-2000;	2000US-0240360.	CC	including cancer. The present sequence is a coding sequence of the
PR	20-OCT-2000;	2000US-0241221.	CC	invention.
PR	20-OCT-2000;	2000US-0241785.	XX	
PR	20-OCT-2000;	2000US-0241786.	XX	Sequence 578 BP; 130 A; 151 C; 166 G; 126 T; 5 other;
PR	20-OCT-2000;	2000US-0241787.		
PR	20-OCT-2000;	2000US-0241808.		
PR	20-OCT-2000;	2000US-0241809.		Query Match 8.9%; Score 484.4; DB 22; Length 578;
PR	20-OCT-2000;	2000US-0241826.		Best Local Similarity 99.4%; Pred. No. 5e-110; Indels 1; Gaps 1;
PR	01-NOV-2000;	2000US-024617.		Matches 496; Conservative 0; Mismatches 2;
PR	08-NOV-2000;	2000US-0246474.	QY	1375 CAAGACCTTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGGCGCTGATTATGTG 1434
PR	08-NOV-2000;	2000US-0246475.	Db	2 CAAGACCTTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGGCGCTGATTATGTG 61
PR	08-NOV-2000;	2000US-0246476.		
PR	08-NOV-2000;	2000US-0246477.		
PR	08-NOV-2000;	2000US-0246478.	QY	1435 GAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACC 1494
PR	08-NOV-2000;	2000US-0246523.	Db	62 GAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACC 121
PR	08-NOV-2000;	2000US-0246524.		
PR	08-NOV-2000;	2000US-0246525.		
PR	08-NOV-2000;	2000US-0246526.	QY	1495 CAAGAACCTCGCATAGTCTACTGTCAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGGCC 1554
PR	08-NOV-2000;	2000US-0246527.	Db	122 CAAGAACCTCGCATAGTCTACTGTCAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGGCC 181
PR	08-NOV-2000;	2000US-0246528.		
PR	08-NOV-2000;	2000US-0246532.	QY	1555 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGACCGCTTCAGCATGTC 1614
PR	08-NOV-2000;	2000US-0246609.	Db	182 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGACCGCTTCAGCATGTC 241
PR	08-NOV-2000;	2000US-0246610.		
PR	08-NOV-2000;	2000US-0246611.		
PR	08-NOV-2000;	2000US-0246613.	QY	1615 TTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGGTCTCGCTGAGCTCATC 1674
PR	17-NOV-2000;	2000US-0249207.	Db	242 TTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGGTCTCGCTGAGCTCATC 301
PR	17-NOV-2000;	2000US-0249208.		
PR	17-NOV-2000;	2000US-0249209.	QY	1675 GGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGG 1734
PR	17-NOV-2000;	2000US-0249210.	Db	302 GGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGG 361
PR	17-NOV-2000;	2000US-0249211.		
PR	17-NOV-2000;	2000US-0249212.	QY	1735 CTGCTCTTCATCCTCGATGGTGTAGATGACAGGATGGTCTTTCAGGAGCGGAGTTCT 1794
PR	17-NOV-2000;	2000US-0249213.	Db	362 CTGCTCTTCATCCTCGATGGTGTAGATGACAGGATGGTCTTTCAGGAGCGGAGTTCT 421
PR	17-NOV-2000;	2000US-0249214.		
PR	17-NOV-2000;	2000US-0249215.		
PR	17-NOV-2000;	2000US-0249216.	QY	1795 GAGCTCTGTCTGCACCTGGAGCCAGCCACAG-CCGGCGGATGCACCTCTGGGCGAGTTTGT 1853
PR	17-NOV-2000;	2000US-0249217.	Db	422 GAGCTCTGTCTGCACCTGGAGCCAGCCACAGCCCGCGGATGCACCTCTGGGCGAGTTTGT 491
PR	17-NOV-2000;	2000US-0249218.		
PR	17-NOV-2000;	2000US-0249219.		
PR	17-NOV-2000;	2000US-0249244.		
PR	17-NOV-2000;	2000US-0249245.		
PR	17-NOV-2000;	2000US-0249246.		
PR	17-NOV-2000;	2000US-0249264.		
PR	17-NOV-2000;	2000US-0249265.		
PR	17-NOV-2000;	2000US-0249297.	QY	1854 GGGGAAAACATACTTCTCC 1872
PR	17-NOV-2000;	2000US-0249299.	Db	482 GGGGAAAACATACTTCTCC 500
PR	17-NOV-2000;	2000US-0249299.		
PR	17-NOV-2000;	2000US-0249300.		
PR	01-DEC-2000;	2000US-0250160.		
PR	05-DEC-2000;	2000US-0250391.		
PR	05-DEC-2000;	2000US-0251030.		
PR	05-DEC-2000;	2000US-0251988.		
PR	05-DEC-2000;	2000US-0256719.		
PR	06-DEC-2000;	2000US-0251479.		
PR	08-DEC-2000;	2000US-0251856.		
PR	08-DEC-2000;	2000US-0251856.		
PR	08-DEC-2000;	2000US-0251868.		
PR	08-DEC-2000;	2000US-0251869.		
PR	08-DEC-2000;	2000US-0251989.		
PR	08-DEC-2000;	2000US-0251990.		
PR	11-DEC-2000;	2000US-0254097.		
PR	05-JAN-2001;	2001US-0259678.		
XX	(HUMA-) HUMAN GENOME SCI INC.			
PA	Rosen CA, Barash SC, Ruben SM;			
XX	WPI; 2001-465570/50.			
XX	P-PSDB; AAM95774.			
XX	Isolated nucleic acid molecule encoding a reproductive system antigen			
PT	is used in preventing, treating or ameliorating a medical condition -			

RESULT 13
ABL97037
ID ABL97037 standard; cDNA; 578 BP.
XX
AC ABL97037;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding cDNA SEQ ID NO: 705.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX

PF 17-JAN-2001; 2001WO-US01329.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235634.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02559678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483232/52.
 XX Nucleic acids encoding 973 human testicular antigen polypeptides,
 DR
 XX
 PT

QY 5104 CTCGGATCCATCCAGGCGCAAGAGGAAATAGGAGGACATGGAACCATTTGGCTCTCG 5163
 Db 127 CTCGGATCCATCCAGGCGCAAGAGGAAATAGGAGGACATGGAACCATTTGGCTCTCG 68
 QY 5164 CTGTGTACAGGCTGAGCCGCCAAAATTGGGGTTTCAGCGTGGAGGCCACGTTGATTCTTG 5223
 Db 67 CTGTGTACAGGCTGAGCCGCCAAAATTGGGGTTTCAGCGTGGAGGCCACGTTGATTCTTG 8
 QY 5224 GCTTTGT 5230
 Db 7 GCTTTGT 1

RESULT 15
 ABV97891/C
 ID ABV97891 standard; cDNA; 416 BP.
 AC ABV97891;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 DE Human pancreatic cancer expressed cDNA SEQ ID NO 3299.
 XX
 KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
 KW cytotstatic; tumour; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200260317-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 30-JAN-2002; 2002WO-US02781.
 XX
 PR 30-JAN-2001; 2001US-265305P.
 PR 31-JAN-2001; 2001US-265682P.
 PR 09-FEB-2001; 2001US-267588P.
 PR 21-MAR-2001; 2001US-278651P.
 PR 28-APR-2001; 2001US-287112P.
 PR 16-MAY-2001; 2001US-291631P.
 PR 12-JUL-2001; 2001US-305484P.
 PR 20-AUG-2001; 2001US-313999P.
 PR 27-NOV-2001; 2001US-333626P.
 XX
 PA (CORI-) CORIXA CORP.

PA Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
 PI WPI; 2002-627435/67.
 XX
 XX New isolated polynucleotide and pancreatic tumor polypeptides, useful
 PT for diagnosing, preventing and/or treating cancer, particularly
 PT pancreatic cancer
 XX
 XX Claim 1; SEQ ID NO 3299; 300pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising: (a)
 CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
 CC (b) complements of (a); (c) sequences consisting of at least 20
 CC contiguous residues of (a); (d) sequences that hybridize to (a), under
 CC moderately stringent conditions; (e) sequences having at least 75% or 90%
 CC identity to (a); or (f) degenerate variants of (a). Polypeptides
 CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
 CC detect cancer in a patient and compositions comprising polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations and
 CC antigen presenting cells expressing the polypeptide are useful in
 CC treating pancreatic cancer and stimulating an immune response. The
 CC polynucleotides can be used as probes or primers for nucleic acid
 CC hybridisation, in the design and preparation of ribozyme molecules for
 CC inhibiting expression of the tumour polypeptides and proteins in the
 CC tumour cells, in vaccines and for gene therapy.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 416 BP; 101 A; 113 C; 114 G; 88 T; 0 other;

Query Match 7.6%; Score 414.4; DB 24; Length 416;
 Best Local Similarity 99.8%; Pred. No. 1.1e-92;
 Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4815 TATCAACACACAGCCCTTGACCCCTTGAGTCTGGCTTTGGCTGACCCCTTTCTTTGGGTCTCA 4874
 Db 416 TATCAACACACAGCCCTTGACCCCTTGAGTCTGGCTTTGGCTGACCCCTTTCTTTGGGTCTCA 357
 QY 4875 GTTCTCTTCTCTGCAACCAAGTTGCCATCTGTTTGGCTTCCAGCACTAAAGTAATGAA 4934
 Db 356 GTTCTCTTCTCTGCAACCAAGTTGCCATCTGTTTGGCTTCCAGCACTAAAGTAATGAA 297
 QY 4935 CTTTGATGATGCTTTTGTGGGCATTTATGTGTCCATGCCAGGGATGCCACAGGGGGCCCC 4994
 Db 296 CTTTGATGATGCTTTTGTGGGCATTTATGTGTCCATGCCAGGGATGCCACAGGGGGCCCC 237
 QY 4995 AGTCCAGGTGGCTTAACAGCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGC 5054
 Db 236 AGTCCAGGTGGCTTAACAGCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGC 177
 QY 5055 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCCAAGCCTAGAGCCCTCCGGATCCC 5114
 Db 176 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCCAAGCCTAGAGCCCTCCGGATCCC 117
 QY 5115 ATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCATTTGCCCTCTGGCTGTGTACAG 5174
 Db 116 ATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCATTTGCCCTCTGGCTGTGTACAG 57
 QY 5175 GGTGAGCCGCCAAAATTGGGGTTTCAGCGTGGAGGCCACGTTGATTCTTTGGCTTTGT 5230
 Db 56 GGTGAGCCGCCAAAATTGGGGTTTCAGCGTGGAGGCCACGTTGATTCTTTGGCTTTGT 1

Search completed: January 29, 2004, 03:29:21
 Job time : 1357.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:01:52 ; Search time 10849.3 Seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggccctggagaggt.....aggaataagaagtacctac 5444

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pin:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	994	18.3	1201	9	AL541851
2	983.2	18.1	1201	13	BX341134
3	898.4	16.5	1201	13	BX341135
4	819.8	15.1	1055	13	BQ072636

ALIGNMENTS

RESULT 1
AL541851
LOCUS
DEFINITION AL541851 Homo sapiens PLACENTA Homo sapiens CDNA clone CS00007D22
5-PRIME, mRNA sequence.
ACCESSION AL541851
VERSION AL541851
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On Feb 15, 2001 this sequence version replaced gi:12873320.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS00007D22&cluster=8396.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

5	814.4	15.0	1046	12	BM549709	BM549709	AGENCOURT
6	787.4	14.5	1110	13	BX457482	BX457482	AGENCOURT
7	767	14.1	768	12	BI518887	BI518887	AGENCOURT
8	749.4	13.8	894	14	CD515485	CD515485	AGENCOURT
9	727	13.4	791	10	BG682080	BG682080	AGENCOURT
10	722.4	13.3	764	13	BQ576345	BQ576345	UI-H-E21-
11	671	12.3	686	14	CA308213	CA308213	UI-H-F11-
12	633.6	11.6	672	12	BI911134	BI911134	603067546
13	633	11.6	728	12	BI908442	BI908442	603067392
14	628.4	11.5	655	12	BQ003173	BQ003173	UI-H-E11-
15	623.6	11.5	615	12	BM152462	BM152462	TCBAP1E90
16	615	11.3	615	12	BM791087	BM791087	K-EST0071
17	575	10.6	592	13	BQ633947	BQ633947	UI-H-F11-
18	566	10.4	566	10	BF059271	BF059271	7K57C12.X
19	558.4	10.3	560	12	BM152839	BM152839	TCBAP1Q12
20	544.4	10.0	554	12	BM193542	BM193542	TCBAP2Q14
21	520.8	9.6	524	9	AW173055	AW173055	XJ82F03.X
22	520.4	9.6	523	12	BM153182	BM153182	TCBAP1Q14
23	509	9.3	595	13	BQ355077	BQ355077	MR2-HT116
24	502.6	9.2	514	9	AV682872	AV682872	AV682872
25	495.4	9.1	580	12	BM152807	BM152807	TCBAP1Q12
26	494	9.1	494	9	AW593727	AW593727	Hg21G12.X
27	489.6	9.0	806	12	BI519523	BI519523	603061807
28	485	8.9	494	12	BM151763	BM151763	TCBAP1E10
29	483.4	8.9	500	10	BE775300	BE775300	QV0-UM009
30	476.4	8.8	633	12	BI909022	BI909022	603070167
31	475	8.7	509	12	BG774227	BG774227	602662231
32	474.6	8.7	521	9	AI819334	AI819334	Wg55h07.X
33	469.8	8.6	504	9	AI742596	AI742596	Wg55h07.X
34	469	8.6	469	9	AI825353	AI825353	Wb17408.X
35	458	8.4	458	10	BE464469	BE464469	hkf17e05.X
36	456.8	8.4	511	13	BX474531	BX474531	DZF2P686G
37	449	8.2	449	9	AI365110	AI365110	qx96b04.X
38	445.4	8.2	461	9	AW291804	AW291804	UI-H-B12-
39	445.2	8.2	487	12	BI028123	BI028123	CM4-MT028
40	441	8.1	464	9	AW291769	AW291769	UI-H-B12-
41	438	8.0	441	9	AI339112	AI339112	qt06b07.X
42	434.4	8.0	466	12	BM193358	BM193358	TCBAP1E96
43	434.2	8.0	791	28	BZ129303	BZ129303	CH230-345
44	424	7.8	425	9	AI439770	AI439770	t161C01.X
45	423.2	7.8	489	12	BM151640	BM151640	TCBAP1E10

QY	4426	TGGAGGCTTGGTGAACACAGAGATCTATGCTGCTGCAACTACTCTGATCCCTCCAGCC	4485
Db	997	TGTGAGSCCTTGTGAACACAGAGATCTATGCTGCTGCAACTACTCTGATCCCTCCAGCC	938
QY	4486	CGCATAGCGTACCTTACCTCTGATGCGCCGCGAGTTGCTGCACTTTGTGACAGTAT	4545
Db	937	CGCATAGCGTACCTTACCTCTGATGCGCCGCGAGTTGCTGCACTTTGTGACAGTAT	878
QY	4546	CGAGAGCAGCTGATAGCCGAGTGAATCGGTGGAGGTTGCTTGGACAACTGATGGA	4605
Db	877	CGAGAGCAGCTGATAGCCGAGTGAATCGGTGGAGGTTGCTTGGACAACTGATGGA	818
QY	4606	CAGGTGCTGAGCCAGAGCAGTACGAGAGGTTGCTGCTGAGAAACAGAGGCCAGCAG	4665
Db	817	CAGGTGCTGAGCCAGAGCAGTACGAGAGGTTGCTGCTGAGAAACAGAGGCCAGCAG	758
QY	4666	ATGCGGAGCTTTCAGCTTGAGCCAGTCTGCGGACCGGAGTGCATAGATGATCTTAC	4725
Db	757	ATGCGGAGCTTTCAGCTTGAGCCAGTCTGCGGACCGGAGTGCATAGATGATCTTAC	698
QY	4726	CAAGCCCTGAAGGAGACCCATCTCACTTATATGGAATCTGGAGAGGCGCAGCAAA	4785
Db	697	CAAGCCCTGAAGGAGACCCATCTCACTTATATGGAATCTGGAGAGGCGCAGCAAA	638
QY	4786	AAGGAGCTCTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCTTGAGTCT	4845
Db	637	AAGGAGCTCTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCTTGAGTCT	578
QY	4846	GGCTTTGGTACCTTCTTTGGTCTGAGTTCTTCTTCTGCAACAGTTGCCATCTG	4905
Db	577	GGCTTTGGTACCTTCTTTGGTCTGAGTTCTTCTTCTGCAACAGTTGCCATCTG	518
QY	4906	GTTCGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTCTGCGGCAATATG	4965
Db	517	GTTCGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTCTGCGGCAATATG	458
QY	4966	TCCATGCCAGGATGCAACAGGGGGCCCGAGTCCAGGTCGCTTAACAGCATCTCAGGAA	5025
Db	457	TCCATGCCAGGATGCAACAGGGGGCCCGAGTCCAGGTCGCTTAACAGCATCTCAGGAA	398
QY	5026	TGTCATCTGGAGCTGGCAAGACCTCTGAGACCTCATAGACCTCATCTGTGGCCACA	5085
Db	397	TGTCATCTGGAGCTGGCAAGACCTCTGAGACCTCATAGACCTCATCTGTGGCCACA	338
QY	5086	GCAGCAAGCTAGAGCCCTCCGATCCCATCCAGCGCAAGAGAGATAGAGGACAT	5145
Db	337	GCAGCAAGCTAGAGCCCTCCGATCCCATCCAGCGCAAGAGAGATAGAGGACAT	278
QY	5146	GGAAACCATTTGCTGCTGTGTCAACAGGCTGAGCCCAAAATTTGGGTTTCAGGTTGG	5205
Db	277	GGAAACCATTTGCTGCTGTGTCAACAGGCTGAGCCCAAAATTTGGGTTTCAGGTTGG	218
QY	5206	AGCCACGTGGATTTTGGCTTTGTACAGAGATCTACAAGAGCAAGCAACAGAGTAA	5265
Db	217	AGCCACGTGGATTTTGGCTTTGTACAGAGATCTACAAGAGCAAGCAACAGAGTAA	158
QY	5266	AGTGAAGGAAGTTTATTCAGAAAATAAAGGAGTATCACAGCTCTTTTGAATTTGCTA	5325
Db	157	AGTGAAGGAAGTTTATTCAGAAAATAAAGGAGTATCACAGCTCTTTTGAATTTGCTA	98
QY	5326	GCAGGCTTTTCCAGTTTATTCAGAAAACCCCTTAAATTTTAAATTTTAAATTT	5385
Db	97	GCAGGCTTTTCCAGTTTATTCAGAAAACCCCTTAAATTTTAAATTTTAAATTT	38
QY	5386	AAGAATTAATAAATAAATAAAGAA	5411
Db	37	AAGAATTAATAAATAAATAAAGAA	12

RESULT 3
BX341135
LOCUS

BX341135 1201 bp mRNA linear EST 01-MAY-2003

DEFINITION BX341135 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ012YC02 5-PRIME, mRNA sequence.
ACCESSION BX341135
VERSION BX341135.1 GI:30310958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ012BB01QP1&cluster=8396.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ012BB01QP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ012YC02"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 297 a 329 c 312 g 217 t 46 others
ORIGIN

Query Match 16.5%; Score 898.4; DB 13; Length 1201;
Best Local Similarity 97.4%; Pred. No. 6.2e-111;
Matches 937; Conservative 7; Mismatches 13; Indels 5; Gaps 3;
QY 773 CCCAGCCCGAGAGGGGCGAGCCACTCTCCCTCTATTCCCTTACAGCCCAAGTGAACCCC 832
Db 57 CCCGGATCAGAGAGGGGCGAGCCACTCTCCCTCTATTCCCTTACAGCCCAAGTGAACCCC 116
QY 833 ACCTGGGTCTCCAGCAACCCACCTCCACCGCAGTGTCTATGCTCCCTGGATCCATGAAT 892
Db 117 ACCTGGGTCTCCAGCAACCCACCTCCACCGCAGTGTCTATGCTCCCTGGATCCATGAAT 176
QY 893 TCCCGGGGGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGTGCCTGACACAT 952
Db 177 TCCCGGGGGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGTGCCTGACACAT 236
QY 953 CTGGAGCGCGCTGGAGAGAAATCTCTGCTCACTCTCTTACCAGCTCTTCCAAAGTCCC 1012
Db 237 CTGGAGCGCGCTGGAGAGAAATCTCTGCTCACTCTCTTACCAGCTCTTCCAAAGTCCC 296
QY 1013 CAGACCATGAGTCTCCAAAGCCAGAGTACCACCGCCGCCACATCCACAGCAGTGTCTGG 1072
Db 297 CAGACCATGAGTCTCCAAAGCCAGAGTACCACCGCCGCCACATCCACAGCAGTGTCTGG 356
QY 1073 GGAGCTGGGGATCCCACTCAGCCCAAGCTAGCAGCCAGAGAGAGGAGGCTCTCTGGGA 1132
Db 357 GGAGCTGGGGATCCCACTCAGCCCAAGCTAGCAGCCAGAGAGGAGGCTCTCTGGGA 416
QY 1133 CCAATGGCTCTGGATGAACGTGAGAAATTTACTACAGAAATCAGAGAAAGAGAGA 1192
Db 417 CCAATGGCTCTGGATGAACGTGAGAAATTTACTACAGAAATCAGAGAAAGAGAGA 476


```

QY 780 CCAGGAAGGGGAGGCCACTCTCCCTCATCTCCCTACAGCCCAAGTGAACCCCACT-GG 838
Db 830 CAGGGAAGGGGAGGCCACTCTCTCTCATCTCCCTAAAGCCCAAGTGACCCCACTGGG 889
QY 839 GGTCTCCCAAGCAACCC 855
Db 890 GGTCTCCCAAGCAACCC 906

RESULT 5
BM549709 1046 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6544160 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745766
DEFINITION 5', mRNA sequence.
ACCESSION BM549709
VERSION BM549709.1 GI:18785339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1046)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12769 row: d column: 23
High quality sequence start: 14
High quality sequence stop: 649.
Location/Qualifiers
1. .1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5745766"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/clone_lib="NIH MGC 118"
/notes="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dr primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invtrorgen). Research Genetics tracking code 027. Note:
this is a NIH MGC Library."
BASE COUNT 227 a 293 c 295 g 231 t
ORIGIN

Query Match 15.0%; Score 814.4; DB 12; Length 1046;
Best Local Similarity 98.4%; Pred. No. 1.1e-99;
Matches 844; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 3712 TCTCAAGGGGAGCTGCATAGAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCCACG 3771
Db 20 TCTCAAGGGGAGCTGCATAGAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCCACG 79

QY 3772 GGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACCTTCCTC 3831
Db 80 GGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACCTTCCTC 139

QY 3832 GTAGCTGGCTCTTACCGCTGGCCCAACACGGGTCTCTGCTTTGTATGATGAGAGAGCGGTG 3891
Db 140 GTAGCTGGCTCTTACCGCTGGCCCAACACGGGTCTCTGCTTTGTATGATGAGAGAGCGGTG 199

```

FEATURES

source

```

3892 ACCGTTGAGATGAATCTGTGTGGGAGCAGTTCTCGGTGAGATCAACCCACAGCAC 3951
200 ACCGTTGAGATGAATCTGTGTGGGAGCAGTTCTCGGTGAGATCAACCCACAGCAC 259
3952 AGCTGGATGTGGCAGGCGCTCTGCTGACATCAAGGCTCAGGCTGAGCCTGGAGCTGTGAAAGCT 4011
260 AGCTGGATGTGGCAGGCGCTCTGCTGACATCAAGGCTCAGGCTGAGCCTGGAGCTGTGAAAGCT 319
4012 GTGACCTCCCTCACTTTGTGGCTCTCCAAGGGGCGCATGTGGACACATCCCTGTTCCAA 4071
320 GTGACCTCCCTCACTTTGTGGCTCTCCAAGGGGCGCATGTGGACACATCCCTGTTCCAA 379
4072 ATGGCCCACTTTAAAGAGAGGAGGTCTCTCGGAGAGCCAGCCAGGCTGAGCTGAT 4131
380 ATGGCCCACTTTAAAGAGAGGAGGTCTCTCGGAGAGCCAGCCAGGCTGAGCTGAT 439
4132 CACATAGTTCTGGAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAATGATCCAT 4191
440 CACATAGTTCTGGAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAATGATCCAT 499
4192 AATGCCCTGGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTAGCACCGCGCTCATCT 4251
500 AATGCCCTGGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTAGCACCGCGCTCATCT 559
4252 GAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCATTTGGGAAGAACTG 4311
560 GAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCATTTGGGAAGAACTG 619
4312 GAGCTCTGATCGAAGCCCTGGAGAACAGCTGTTCTCGGAGTCTACCTTGGCCAC 4371
620 GAGCTCTGATCGAAGCCCTGGAGAACAGCTGTTCTCGGAGTCTACCTTGGCCAC 679
4372 TTGGGATCAGGATCAGCTGCAAGTGAAGAAAGAGATGAGACTCTGTG-TGTGGA 4430
680 TTGGGATCAGGATCAGCTGCAAGTGAAGAAAGAGATGAGACTCTGTG-TGTGGA 739
4431 GGCCTTGTGAAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCGCGAT 4490
740 GGCCTTGTGAAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCGCGAT 799
4491 AGCCGTACTCTTACCTCTGGATGCCCGGAGTTGC--TGCACCTTTGTGAGCAGTATCGA 4548
800 AGCCGTACTCTTACCTCTGGATGCCCGGAGTTGCCTGCACTTTGGGGAGCAGTATCGA 859
4549 GAGCAGCTGATAGCCCGA 4566
860 GAGCAGCTGATAGCCCGA 877

RESULT 6
BX457482 1110 bp mRNA linear EST 22-MAY-2003
LOCUS BX457482 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YN15
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX457482
VERSION BX457482.1 GI:31028889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1110)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see

```

```

http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008CG08QPl&cluster=8396.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0CAP008CG08QPl.
Location/Qualifiers
1. .1110
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008QPl5"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

FEATURES

source

BASE COUNT	218 a	270 c	257 g	262 t	103 others
ORIGIN	Library was not normalized.				
Query Match	14.5%; Score 787.4; DB 13; Length 1110;				
Best Local Similarity	88.3%; Pred. No. 4.6e-96;				
Matches 848; Conservative	58; Mismatches 48; Indels 6; Gaps 6;				
QY	1644	GTCCAAAGTGGTGAGTCTCGCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCC	1703		
DB	72	GACCCACGCGTCCGGTCTCGCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCC	131		
QY	1704	CATTAGACAGATCCTGTCCTTAGCCAGAGCGGCTCTTTCATCTCGATGGTGTAGATGA	1763		
DB	132	CATTAGACAGATCCTGTCCTTAGCCAGAGCGGCTCTTTCATCTCGATGGTGTAGATGA	191		
QY	1764	GCCAGGATGGTCTTTCAGAGCGCGAGTCTGAGCTCTGTCGACATGGAGCCAGCCACA	1823		
DB	192	GCCAGGATGGTCTTTCAGAGCGCGAGTCTGAGCTCTGTCGACATGGAGCCAGCCACA	251		
QY	1824	GCCGGCGGATGCACTGCTGGGCGAGTTTGGCTGGGAAAAACTATATCTCCGAGGCATCCTT	1883		
DB	252	GCCGGCGGATGCACTGCTGGGCGAGTTTGGCTGGGAAAAACTATATCTCCGAGGCATCCTT	311		
QY	1884	CCTGATCA CGGCTCGGACCA CAGCTCTGCGAGAACCTCATCTCTTCTTTTGAGCAGGCACG	1943		
DB	312	CCTGATCACGGCTCGGACCA CAGCTCTGCGAGAACCTCATCTCTTCTTTTGAGCAGGCACG	371		
QY	1944	TTGGGTAGAGTCTCTGGGGTTCTCTGAGTCCAGCAGGAAGGAATTTCTACAGATATTT	2003		
DB	372	TTGGGTAGAGTCTCTGGGGTTCTCTGAGTCCAGCAGGAAGGAATTTCTACAGATATTT	431		
QY	2004	CACAGATGAAGCGCAAGCAATTAGAGCCTTTTAGTTGGTCAAAACAAGAGCTCTG	2063		
DB	432	CACAGATGAAGCGCAAGCAATTAGAGCCTTTAGTTGGTCAAAACAAGAGCTCTG	491		
QY	2064	GGCCCTGTGTTGTGCCCTGGGTGTCTGGCTGGCCCTGCACTTGGCTGATGCGACGAGAT	2123		
DB	492	GGCCCTGTGTTGTGCCCTGGGTGTCTGGCTGGCCCTGCACTTGGCTGATGCGACGAGAT	551		
QY	2124	GAAGCGGAAGAAAAACTCACATGACTTCCAGAGACCACACACACCTCTGTCTACATTA	2183		
DB	552	GAAGCGGAAGAAAAACTCACATGACTTCCAGAGACCACACACACCTCTGTCTACATTA	611		
QY	2184	CTTTGCCCGGCTCTCCAAGCTCAGCATTTGGGACCCAGCTCAGAGA - CCTCTGCTCTC	2242		
DB	612	CTTTGCCCGGCTCTCCAAGCTCAGCATTTGGGACCCAGCTCAGAGCCCTCTGCTCTC	671		
QY	2243	TGGCTGCTGAGGGCATCTGGCAAAAA - AAGACCCCTTTTTCAGTCAGATGACCTCAGGAAG	2301		
DB	672	TGGCTGCTGAGGGCATCTGGCAAAAAAAGACCCCTTTTTCAGTCAGATGACCTCAGGAAG	731		
QY	2302	CATGGG - TTAGATGGGGCCATCATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGACGA	2360		
DB	732	CATGGGTTTATAGTGGGSCATCATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGACGA	791		

Qy	2361	CCCCATCCCTCTGAGCTACAGCTTCATTACCTCTGTGTTCCAAAGATT-CTTTGCGACAA	241
Db	792	CCCCATCCCTCTGAGCTACAGCTTCATTACCTCTGTGTTCCAAAGATT-CTTTGCGACAA	851
Qy	2420	TGTCCTATGCTCTGGAGGATGAGAAGGGGAGAGAGTAAACATCTTAATTGCATCATAGA-T	2478
Db	852	TGTCCTATGCTCTGGAGGATGAGAAGGGGAGAGAGTAAACATCTTAATTGCATCATAGA-T	911
Qy	2479	TTGGAAAAGACGGCTAGA-AGCATATGGAATACATGGCTGTTTGGGCGATCAACACACG	2537
Db	912	TTGGAAAAGACGGCTAGA-AGCATATGGAATACATGGCTGTTTGGGCGATCAACACACG	971
Qy	2538	TTTTCCTATTGGGGCTGTAAAGTCATGAGGGGGGAGAGAGATGGAGAACATCTTCTCACTG	2597
Db	972	TTTTCCTATTGGGGCTGTAAAGTCATGAGGGGGGAGAGAGATGGAGAACATCTTCTCACTG	1031
RESULT 7			
B1518887			
LOCUS	603061807F1 NIH_MGC_118 Homo sapiens	768 bp mRNA linear	EST 29-AUG-2001
DEFINITION	cDNA clone IMAGE:5211072 5', mRNA sequence.		
ACCESSION	B1518887		
VERSION	B1518887		
KEYWORDS	B1518887.1 GI:15343679		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 768)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ ;		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-1@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M11529 row: n column: 01 High quality sequence stop: 766.		

FEATURES: source

```

1. 768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5211072"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH MGC 118"
/note="Vector; C1W-SPO26;
(destroyed); RNA source leu
non-activated adult donors
and directionally cloned (
cloning). Average insert si
1.2-3.3 Kb. Library is nor
full-length clones and was
(invitrogen). Research Gen
this is a NIH MGC Library."
176 a 213 c 225 g 154 t
BASE COUNT

```

	BASE COUNT	ORIGIN
1	0.0000	0.0000
2	0.0000	0.0000
3	0.0000	0.0000
4	0.0000	0.0000
5	0.0000	0.0000
6	0.0000	0.0000
7	0.0000	0.0000
8	0.0000	0.0000
9	0.0000	0.0000
10	0.0000	0.0000
11	0.0000	0.0000
12	0.0000	0.0000
13	0.0000	0.0000
14	0.0000	0.0000
15	0.0000	0.0000
16	0.0000	0.0000
17	0.0000	0.0000
18	0.0000	0.0000
19	0.0000	0.0000
20	0.0000	0.0000
21	0.0000	0.0000
22	0.0000	0.0000
23	0.0000	0.0000
24	0.0000	0.0000
25	0.0000	0.0000
26	0.0000	0.0000
27	0.0000	0.0000
28	0.0000	0.0000
29	0.0000	0.0000
30	0.0000	0.0000
31	0.0000	0.0000
32	0.0000	0.0000
33	0.0000	0.0000
34	0.0000	0.0000
35	0.0000	0.0000
36	0.0000	0.0000
37	0.0000	0.0000
38	0.0000	0.0000
39	0.0000	0.0000
40	0.0000	0.0000
41	0.0000	0.0000
42	0.0000	0.0000
43	0.0000	0.0000
44	0.0000	0.0000
45	0.0000	0.0000
46	0.0000	0.0000
47	0.0000	0.0000
48	0.0000	0.0000
49	0.0000	0.0000
50	0.0000	0.0000
51	0.0000	0.0000
52	0.0000	0.0000
53	0.0000	0.0000
54	0.0000	0.0000
55	0.0000	0.0000
56	0.0000	0.0000
57	0.0000	0.0000
58	0.0000	0.0000
59	0.0000	0.0000
60	0.0000	0.0000
61	0.0000	0.0000
62	0.0000	0.0000
63	0.0000	0.0000
64	0.0000	0.0000
65	0.0000	0.0000
66	0.0000	0.0000
67	0.0000	0.0000
68	0.0000	0.0000
69	0.0000	0.0000
70	0.0000	0.0000
71	0.0000	0.0000
72	0.0000	0.0000
73	0.0000	0.0000
74	0.0000	0.0000
75	0.0000	0.0000
76	0.0000	0.0000
77	0.0000	0.0000
78	0.0000	0.0000
79	0.0000	0.0000
80	0.0000	0.0000
81	0.0000	0.0000
82	0.0000	0.0000
83	0.0000	0.0000
84	0.0000	0.0000
85	0.0000	0.0000
86	0.0000	0.0000
87	0.0000	0.0000
88	0.0000	0.0000
89	0.0000	0.0000
90	0.0000	0.0000
91	0.0000	0.0000
92	0.0000	0.0000
93	0.0000	0.0000
94	0.0000	0.0000
95	0.0000	0.0000
96	0.0000	0.0000
97	0.0000	0.0000
98	0.0000	0.0000
99	0.0000	0.0000
100	0.0000	0.0000

Query Match	14.1%	Score 767;	DB 12;	Length 768;
Best Local Similarity	100.0%;	Prod. No. 2.7e-93;		
Matches 767;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3119	CCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTT	3178	
Db	1	CCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTT	60	


```

Db      661  |||||CTGCTTGTGATGAGAGAGCGGTGACCGTTGAGATTCAATTCCTGTGTGGGAC 720
Qy      3922 CAGTT--CCTGGGTGAGATCAACCCACACACAGCTGGATGGTGGCAGGCGCTCTGCTGG 3979
Db      721  CAGTTTCTCGGGTGGATGATCAACCCACACAGACANCTGGGATGGTGGCAGGCGCTCTTGC 780
Qy      3980 ACATCAAGGCTGAGCGCTGGAGCTGTGGAAGCTGTG 4014
Db      781  TGGNACATCAAGGCTTGAACCTGGGAAGCTGTG 815

RESULT 9
BG682080
LOCUS
DEFINITION BG682080.1 GI:13913477
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10616 row: b column: 16
High quality sequence stop: 770.
Location/Qualifiers
1. .791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754607"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 195 a 203 c 200 g 193 t
ORIGIN

Query Match 13.4%; Score 727; DB 10; Length 791;
Best Local Similarity 98.2%; Pred. No. 6.1e-88;
Matches 777; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

Qy      2045 AATCAACAAAGAGCTCTGGCGCCCTGTCTGTGTGCGCTGGGTGCTCGCTGGCTGCA 2104
Db      3 AATCACAAAGAGCTCTGGCGCCCTGTCTGTGTGCGCTGGGTGCTCGCTGGCTGCA 61
Qy      2105 CTGCTGTGATGACGATGAGCGGAAGAAAACCTACACTGACTTCCAAAGACACCA 2164
Db      62 CTGCTGTGATGACGATGAGCGGAAGAAAACCTACACTGACTTCCAAAGACACCA 121
Qy      2165 CAACCTCTGTCTACATTACCTTGGCCAGCTCTCCAAAGCTCAGCATTTGGACCCCGC 2224
Db      122 CAACCTCTGTCTACATTACCTTGGCCAGCTCTCCAAAGCTCAGCATTTGGACCCCGC 181
Qy      2225 TCAGAGACCTCTGCTCTCTGCTGTGAGGCGCATCTGGCAAAAAAGACCCCTTTTCAGTC 2284

```

```

Db      182 TCAGAGACCTCTGCTCTCTGCTGCTGAGGGCATCTGCAAAAAAGACCCCTTTTCAGTC 241
Qy      2285 CAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTTCTTTGAAGATGG 2344
Db      242 CAGATGACCTCAGGAGCATGGTTAGATGGGGCCATCATCTCCACCTTCTTTGAAGATGG 301
Qy      2345 GTATTCTTCAAGAGACACCCCATCCCTCTGAGCTACAGCTTCATTCAACCTCTGTTTCCAAG 2404
Db      302 GTATTCTTCAAGAGACACCCCATCCCTCTGAGCTACAGCTTCATTCAACCTCTGTTTCCAAG 361
Qy      2405 AGTTCTTTGAGCAATGTCCTATGCTTGGAGGATGAGAGGGGAGAGGTAACATTCTA 2464
Db      362 AGTTCTTTGAGCAATGTCCTATGCTTGGAGGATGAGAGGGGAGAGGTAACATTCTA 421
Qy      2465 ATTGCATCATAGATTGTTGAAAAGACGCTAGAGACATATGGAATACATGCGCTGTTTGGGG 2524
Db      422 ATTGCATCATAGATTGTTGAAAAGACGCTAGAGACATATGGAATACATGCGCTGTTTGGGG 481
Qy      2525 CATCAACACACGTTTCTCTATTGGGCTGTTTAAAGTGAATGAGGGGAGAGAGATGGAGA 2584
Db      482 CATCAACACACGTTTCTCTATTGGGCTGTTTAAAGTGAATGAGGGGAGAGAGATGGAGA 541
Qy      2585 ACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGACAGTGGGTCCGCTCCCTGC 2644
Db      542 ACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACCTGATGACAGTGGGTCCGCTCCCTGC 601
Qy      2645 AGCTGCTGCTGACGCCACACTCTCTGGAGTCCCTCCACTGCTGTACAGAGACTCGGAACA 2704
Db      602 AGCTGCTGCTGACGCCACACTCTCTGGAGTCCCTCCACTGCTGTACAGAGACTCGGAACA 661
Qy      2705 AAACGTTCTCTGACACAGTGAATGCGCCATTTTGAAGAAATGGGCA-TGTGTGTAGAAACA 2763
Db      662 AAA-GTTCTCTGACACAGTGAATGCGCCATTTTGAAGAAATGGGCA-TGTGTGTAGAAACA 720
Qy      2764 GACATGAGAGCTCTTAGTGTGACATTTCTGCAATT-AAATTACGCCGACAGTGAAGAGCT 2822
Db      721 GACATGAGAGCTCTTAGTGTGACATTTCTGCAATTAAATTCAGCCGCCACGTTGAAGAGCT 780
Qy      2823 TCAGCTGATTG 2833
Db      781 TCAGCTGATTG 791

RESULT 10
BG576345/c
LOCUS
DEFINITION UI-H-E21-bbi-n-10-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
UI-H-E21-bbi-n-10-0-UI 3', mRNA sequence.
ACCESSION BG576345
VERSION BG576345.1 GI:21479662
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 764)
AUTHORS NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
Orthopedics
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-71, >POLY_A#Simple repeat (matched complement) 72-100,
>AT rich#low complexity 157-242, >MER41B#LTR/MER4-group
Seq primer: M13 FORWARD

```

POLYA=Yes.		Location/Qualifiers		Db	
FEATURES		1. .764		215	TACAGAGCAAGCCACAGAGTAAGTCGAGGAAGTTTATTTCAGAAAATAAGAGTAT 156
source		/organism="Homo sapiens"		5302	CACAGCTCTTTAGAAATTTGTCTAGCAGGCTTCCAGCTTTTCCAGTATTTTACAGAAAAACCCCTATAA 5361
		/mol_type="mRNA"		155	CACAGCTCTTTAGAAATTTGTCTAGCAGGCTTCCAGTATTTTACAGAAAAACCCCTATAA 96
		/db_xref="taxon:9606"		5362	ATTAAAAATTTTCTACTTAAATTTTAAAGATTTAAAAAATACAAAAAGAAAAATGAAAA 5421
		/clone="UI-H-E21-bib-1-n-10-0-UI"		95	ATTAAAAATTTTCTACTTAAATTTTAAAGATTTAAAAAATACAAAAAGAAAAATGAAAA 36
		/tissue_type="Chondrosarcoma Grade II"		5422	TAAAGGAATAAGAGCTTA 5439
		/dev_stage="Adult"		35	TAAAGGAATAAGAGCTTA 18
		/lab_host="DH10B (Life Technologies)"		RESULT 11	
		/clone_lib="NCI CGAP Ch2"		CA308213/c	
		/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.		CA308213	686 bp mRNA linear EST 01-NOV-2002
		TAG LIB=UI-H-E21		UI-H-F11-bib-1-22-0-UI.s1	NCI CGAP_F11 Homo sapiens cDNA clone
		TAG TISSUE=grade-2-chondrosarcoma		UI-H-F11-bib-1-22-0-UI 3'	mRNA sequence.
		TAG SEQ=ATCTAATATG		CA308213	
BASE COUNT	173 a 175 c 175 g 239 t 2 others			CA308213.1	GI:24471267
ORIGIN				EST	
Query Match	13.3%; Score 722.4; DB 13; Length 764;			Homo sapiens (human)	
Best Local Similarity	99.5%; Pred. No. 2.5e-87;			Homo sapiens	
Matches	734; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
QY	4702 CGAAGTGCAAGATGACTCTACCAAGCCCTGAAGAGACCATCTCACCCTCATTATG 4761			NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.
Db	754 CCAGAGTGCAAGAT-GACTCTACCAAGCCCTGAAGAGACCATCTCACCCTCATTATG 696			National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
QY	4762 GAACCTCTGGAGAGGCGCAGCAAAAGGACTCTCCACTCAGCAGCTGAAGTATCAAC 4821			Tumor Gene Index	
Db	695 GAACCTCTGGAGAGGCGCAGCAAAAGGACTCTCCACTCAGCAGCTGAAGTATCAAC 636			Unpublished	
QY	4822 ACCAGCCCTTGACCCCTTGCTGCTTGGCTGACCCCTCTTTGGGTCTCAGTTTCTT 4881			Contact: Robert Strausberg, Ph.D.	
Db	635 ACCAGCCCTTGACCCCTTGCTGCTTGGCTGACCCCTCTTTGGGTCTCAGTTTCTT 576			Email: cgapbs@mail.nih.gov	
QY	4882 TCTCTGCAACAAGTTGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAACTTTGAT 4941			Tissue Procurement: Dr. Gary W. Hunninghake, U of I	
Db	575 TCTCTGCAACAAGTTGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAACTTTGAT 516			cDNA library preparation: Dr. M. Bento Soares, University of Iowa	
QY	4942 GATGCTTTGCTGGCATTTATGTCCATCCAGGGATGCCAGGGGCCCCCAGTCCAG 5001			DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
Db	515 GATGCTTTGCTGGCATTTATGTCCATCCAGGGATGCCAGGGGCCCCCAGTCCAG 456			Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bent-soares@uiowa.edu	
QY	5002 GTGGCCCTAACAGCATCTCAGGAATGTCCATCTGGAGCTGGCAAGACCCCTGCAGACCTC 5061			The following repetitive elements were found in this cDNA sequence: 32-117, >MER41B#LTR/MER4-group	
Db	455 GTGGCCCTAACAGCATCTCAGGAATGTCCATCTGGAGCTGGCAAGACCCCTGCAGACCTC 396			Seq primer: M13 FORWARD	
QY	5062 ATAGAGCCTCATCTGGTGGCCACAGACGCCATAGACCTCCGGATGCCATCCAGG 5121			POLYA=Yes.	
Db	395 ATAGAGCCTCATCTGGTGGCCACAGACGCCATAGACCTCCGGATGCCATCCAGG 336			Location/Qualifiers	
QY	5122 CGCAAGAGGAATAGGAGGACATGGACCATTTGCTCTGGCTGTGTACAGGGTGAGC 5181			1. .686	
Db	335 CGCAAGAGGAATAGGAGGACATGGACCATTTGCTCTGGCTGTGTACAGGGTGAGC 276			/organism="Homo sapiens"	
QY	5182 CCCAAATTTGGGTTTACGCTGGGAGGCGCAGCTGGATTCTTGGCTTTGTACAGCAAGATC 5241			/mol_type="mRNA"	
Db	275 CCCAAATTTGGGTTTACGCTGGGAGGCGCAGCTGGATTCTTGGCTTTGTACAGCAAGATC 216			/db_xref="taxon:9606"	
QY	5242 TACAAGACGACGCAACAGAGTAAAGTGAAGGAGTTTATTTCAGAAAATAAGAGTAT 5301			/clone="UI-H-F11-bib-1-22-0-UI"	
Db				/tissue_type="Aveolar Macrophage"	
				/dev_stage="Adult"	
				/lab_host="DH10B (Life Technologies)"	
				/clone_lib="NCI CGAP_F11"	
				/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_F11 is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCCATGCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.	
				TAG LIB=UI-H-F11	
				TAG TISSUE=Human Lung Aveolar Macrophage	
				TAG_SEQ=GGCCATGCGG	

BASE COUNT	143 a	180 c	174 g	188 t	1 others
ORIGIN					
Query Match	12.3%; Score 671; DB 14; Length 686;				
Best Local Similarity	99.9%; Pred. No. 2e-80;				
Matches 671; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
QY	4646	AGAACACGAGGCCACGACGATCGGGAAGCTGTTTCAGCTTGAGCCAGTCTCTGGACCGGA	4705		
Db	686	AGAACACGAGGCCACGACGATCGGGAAGCTGTTTCAGCTTGAGCCAGTCTCTGNACCGGA	627		
QY	4706	AGTGCAAAAGATGGAATCTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATTTATGGAAC	4765		
Db	626	AGTGCAAAAGATGGAATCTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATTTATGGAAC	567		
QY	4766	TCCTGGGAGAGGCGACGACAAAAGGAGCTCTCTGCCACTCAGCAGCTGAAGTATCAACACCA	4825		
Db	566	TCCTGGGAGAGGCGACGACAAAAGGAGCTCTCTGCCACTCAGCAGCTGAAGTATCAACACCA	507		
QY	4826	GCCTTTGACCTTTGAGTCTCTGGCTTTTGGCTGACCTTCTTTTGGGTCTCAGTTCTTTCTC	4885		
Db	506	GCCTTTGACCTTTGAGTCTCTGGCTTTTGGCTGACCTTCTTTTGGGTCTCAGTTCTTTCTC	447		
QY	4886	TGCAAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACAATAAGTAATGGAACCTTTGATGATG	4945		
Db	446	TGCAAAACAAGTTGCCATCTGGTTTGCCTTCCAGCACAATAAGTAATGGAACCTTTGATGATG	387		
QY	4946	CTTTTGTCTGGGCATTTATGTCTCATCCAGGATGCCACAGGGGGCCAGTCCAGGTGG	5005		
Db	386	CTTTTGTCTGGGCATTTATGTCTCATCCAGGATGCCACAGGGGGCCAGTCCAGGTGG	327		
QY	5006	CCTAACAGCATCTCAGGGAATGTCCATCTGGAGTGGCAAGCCCTGCGACACCTCATAG	5065		
Db	326	CCTAACAGCATCTCAGGGAATGTCCATCTGGAGTGGCAAGCCCTGCGACACCTCATAG	267		
QY	5066	AGCCTCATCTGTTGGCCACAGCAGCAAGCCTTAGAGCCCTCCGGATCCCATCCAGGCGCA	5125		
Db	266	AGCCTCATCTGTTGGCCACAGCAGCAAGCCTTAGAGCCCTCCGGATCCCATCCAGGCGCA	207		
QY	5126	AAGAGGAATAGGAGGACATGGAAACCATTTGCCCTCTGGCTGTGTACAGGGTGAGCCCCA	5185		
Db	206	AAGAGGAATAGGAGGACATGGAAACCATTTGCCCTCTGGCTGTGTACAGGGTGAGCCCCA	147		
QY	5186	AAATTGGGGTTTACGCTGGGAGGCCACGTGGATTCTTTGGCTTTGTACAGGAAGATCTACA	5245		
Db	146	AAATTGGGGTTTACGCTGGGAGGCCACGTGGATTCTTTGGCTTTGTACAGGAAGATCTACA	87		
QY	5246	AGAGCAAGCCACAGAGTAAAGTGAAGGAAGTTTATTTCAGAAATAAAGGAGTATCACA	5305		
Db	86	AGAGCAAGCCACAGAGTAAAGTGAAGGAAGTTTATTTCAGAAATAAAGGAGTATCACA	27		
QY	5306	GCTCTTTTAGAA 5317			
Db	26	GCTCTTTTAGAA 15			
RESULT 12					
BI911134					
LOCUS	603067546F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216770 5',				
DEFINITION	mRNA sequence.				
ACCESSION	BI911134				
VERSION	BI911134.1 GI:16174696				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 672)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished				
	Contact: Robert Strausberg, Ph.D.				

FEATURES					
source	1. 672				
location/Qualifiers					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="IMAGE:5216770"					
/tissue_type="leukocyte"					
/lab_host="DH10B"					
/clone_lib="NIH_MGC_118"					
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned. (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."					
BASE COUNT	138 a	187 c	189 g	158 t	
ORIGIN					
Query Match	11.6%; Score 633.6; DB 12; Length 672;				
Best Local Similarity	99.1%; Pred. No. 2e-75;				
Matches 658; Conservative	0; Mismatches 4; Indels 2; Gaps 2;				
QY	3583	GCTTCCCATGTTGCTCAGGCTAAATCTCAAATCTCTGGAAGTGTGAGCAAGATCTTCCCAATT	3642		
Db	1	GCTTCCCATGTTGCTCAGGCTAAATCTCAAATCTCTGGAAGTGTGAGCAAGATCTTCCCAATT	60		
QY	3643	GCTGAGATTGAGAGGAAGCTCCCGAGAGTAGTACCGGTGGACCTCTCTGGTGGCTCT	3702		
Db	61	GCTGAGATTGAGAGGAAGCTCCCGAGAGTAGTACCGGTGGACCTCTCTGGTGGCTCT	120		
QY	3703	TCTCCT-CCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTG	3761		
Db	121	TCTCCTGGCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTG	180		
QY	3762	GGGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGT	3821		
Db	181	GGGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGT	240		
QY	3822	TCACTTCCCTGTAGTGGCTCTTACCGCTGCCCCAACACGGGTCTCTGCTTTGTGATGAG	3881		
Db	241	TCACTTCCCTGTAGTGGCTCTTACCGCTGCCCCAACACGGGTCTCTGCTTTGTGATGAG	300		
QY	3882	AGAAGCGGTGACCGTTGAGATTGAATCTGTGTGGGACCCAGTTCTCTGGGTGAGATCAA	3941		
Db	301	AGAAGCGGTGACCGTTGAGATTGAATCTGTGTGGGACCCAGTTCTCTGGGTGAGATCAA	360		
QY	3942	CCCAACAGCAGCTGGATGTGGGAGGCGCTCTCTGTGACATCAAGGCTGAGCCCTGGAGC	4001		
Db	361	CCCAACAGCAGCTGGATGTGGGAGGCGCTCTCTGTGACATCAAGGCTGAGCCCTGGAGC	420		
QY	4002	TGTGGAAGCTGACACCTCCCTCTCTGCTCTCCAGGGGGCCATGTGGACATC	4061		
Db	421	TGTGGAAGCTGACACCTCCCTCTCTGCTCTCCAGGGGGCCATGTGGACATC	480		
QY	4062	CCTGTTCCAAATGGCCCCACTTTAAAGAGAGGAGGATGCTCTCTGGAGAGCCAGCAGGT	4121		
Db	481	CCTGTTCCAAATGGCCCCACTTTAAAGAGAGGAGGATGCTCTCTGGAGAGCCAGCAGGT	539		
QY	4122	GGAGTGCATCAGTGTCTGGAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTCTGAA	4181		

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11544 row: k column: 11

High quality sequence stop: 670.

Db 540 GGAGCTGCATCATAGATTCTTGGAAAAACCCAGAGTTCTCCCTTGGGAGTCTCTCTGAA 599

Qy 4182 AATGATCCATAATGCCCTCGGCTTCATTCCTCCGTCACCTCTGTGGTGTGCTTTACACAGC 4241

Db 600 AATGATCCATAATGCCCTCGGCTTCATTCCTCCGTCACCTCTGTGGTGTGCTTTACACAGC 659

Qy 4242 CGTC 4245

Db 660 CGTC 663

RESULT 13

BI908442

LOCUS

DEFINITION 728 bp mRNA linear EST 16-OCT-2001

603067392F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216327 5',

RNA sequence.

ACCESSION BI908442

VERSION BI908442

KEYWORDS EST.

SOURCE BI908442.1 GI:16171396

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11543 row: h column: 24

High quality sequence stop: 718.

FEATURES

Location/Qualifiers

1..728

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5216327"

/tissue_type="leukocyte"

/lab_host="DH10B"

/clone_lib="NIH_MGC_118"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV

(destroyed); RNA source leukocytes from anonymous pool of

non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.7 kb, insert size range

1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 027. Note:

this is a NIH MGC Library."

BASE COUNT 200 a 208 c 200 g 120 t

ORIGIN

Query Match 11.6%; Score 633; DB 12; Length 728;

Best Local Similarity 98.3%; Pred. No. 2.4e-75;

Matches 713; Conservative 0; Mismatches 5; Indels 7; Gaps 7;

Qy 885 CCATGAATTCGGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTGGACAGCTGCC 944

Db 6 CCATGAATTCGGCGGGTGCACCCA-GGCTCAGAGAGAGGGTTTGGACAGCTGCC 64

Qy 945 TGACACATCTGGACGGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTCC 1004

Db 65 TGACACATCTGGACGGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTCC 124

Qy 1005 AAGCTCCCCAGACCATGAGTCTCCAGCCAGGAGTCCACCAACGCCCCACATCCACAGC 1064

Db 125 AAGCTCCCCAGACCATGAGTCTCCAGCCAGGAGTCAACCCGCCCATCCATCCACAGC 184

Qy 1065 AGTGTGGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGAGAGGC 1124

Db 185 AGTGTGGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGAGAGGC 244

Qy 1125 TCCTGGGACCAATGGCTCTGGATGAAACGTGAGGAATTTACTACACAGAAATCAGAGA 1184

Db 245 TCCTGGGACCAATGGCTCTGGATGAAACGTGAGGAATTTACTACACAGAAATCAGAGA 304

Qy 1185 AAGAGAGAGAGAGAAATCA-GAGAAAGGACGCCCCCATGGCAGCGGTGGTAGGAACGC 1243

Db 305 AAGAGAGAGAGAGAAATCAGAGAAAGGACGCCCCCATGGCAGCGGTGGTAGGAACGC 364

Qy 1244 CCCCACAGCGGCACACACAGCCTACAGCCGCCACCCACCCATGGGAGCTCTCTGTGAGAG 1303

Db 365 CCCCACAGCGGCACACACAGCCTACAGCCGCCACCCACCCATGGGAGCTCTCTGTGAGAG 424

Qy 1304 AAGCCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGC 1363

Db 425 AAGCCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGC 484

Qy 1364 TGCTACTTCTCAAGAGACCTCACCCAGAGCCAGATCCCTGTGTCAGAGAACTGSC 1423

Db 485 TGCTACTTCTCAAGAGATCTCACCCAGAGCCAGATCCCTGTGTCAGAGAACTGSC 544

Qy 1424 CTGA-TTATGTGGA-GGAGAAATCGAGGACATTTAATTCAGATCAGAGACTTATTTGGGCC 1481

Db 545 CTGACTTATGTGGCGGAGATCGAGACATTTAATTCAGATCAGAGACTTATTTGG-CC 603

Qy 1482 AGGCTTGATACCCAGAACCTCGCATAGTCATCTGAGGGGGTGTCTGGAATTTGGAA 1541

Db 604 AGGCTTGATACCCAGAACCTCGCATAGTCATCTGAGGGGGTGTCTGGAATTTGGAA 663

Qy 1542 GTCAACACTGGCCAGG-CAGGTGAGGAAGCTGGGGAGAGCCAGCTGTATGGGACC 1600

Db 664 GTCAACACTGGCCAGGCGAGGTGAAGGAAGCTGGGGAGAGG-CAGCTGTATGGGACA 722

Qy 1601 GCTTC 1605

Db 723 GTTC 727

RESULT 14

BQ003173/c

LOCUS

DEFINITION 655 bp mRNA linear EST 26-MAR-2002

UI-H-E11-ayx-n-01-0-UI.s1 NCI_CGAP_E11 Homo sapiens cDNA clone

IMAGE:5845056 3', mRNA sequence.

ACCESSION BQ003173

VERSION BQ003173

KEYWORDS BQ003173.1 GI:19728073

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

The following repetitive elements were found in this cDNA

sequence: 32-117, >MER41B#LTR/MER4-group

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

FEATURES

```

source
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP E11"
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP E11 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(3')18 tail. The sequence tag for this library is
ACACTGGCAC.
TAG LIB=UI-H-E11
TAG TISSUE=chondrosarcoma
TAG_SEQ=ACACTGGCAC"
BASE COUNT 143 a 170 c 164 g 178 t
ORIGIN

Query Match 11.5%; Score 628.4; DB 12; Length 655;
Best Local Similarity 99.7%; Pred. No. 1e-74;
Matches 640; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4676 TGTTGAGCTTACGAGTCTGGGACCGAAGTGAAGATGACTCTACCAAGCCCTGA 4735
DB |||||||
655 TTTTACGCTTGAGCCAGTCTGGGACCGAAGTGAAGATGACTCTACCAAGCCCTGA 596
QY 4736 AGGAGACCCATCTCACTCATTTATGGAAGTCTGGGAGAGGCGCAAAAAGGGACTCC 4795
DB |||||||
595 AGAGAGACCCATCTCACTCATTTATGGAAGTCTGGGAGAGGCGCAAAAAGGGACTCC 536
QY 4796 TGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTTGACCTTTGAGTCTGCGCTTTGGCT 4855
DB |||||||
535 TGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTTGACCTTTGAGTCTGCGCTTTGGCT 476
QY 4856 GACCTTCTTTGGGCTCAGTTTCTTCTGTGCAACAAAGTTGCCATCTGGTTTGCCTTC 4915
DB |||||||
475 GA-CCTTCTTTGGGCTCAGTTTCTTCTGTGCAACAAAGTTGCCATCTGGTTTGCCTTC 417
QY 4916 CAGCAGCTAAAGTAAATGGAAGTCTTGTGATGATGCTTGTGCGGATTAATGTCATGCCAG 4975
DB |||||||
416 CAGCAGCTAAAGTAAATGGAAGTCTTGTGATGATGCTTGTGCGGATTAATGTCATGCCAG 357
QY 4976 GGATGCCACAGGGGGCCCCAGTCCAGTGGCCCTTAACAGCATCTCAGGGAATGTCCATCTG 5035
DB |||||||
356 GGATGCCACAGGGGGCCCCAGTCCAGTGGCCCTTAACAGCATCTCAGGGAATGTCCATCTG 297
QY 5036 GAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGGCCACAGAGCCCAAGC 5095
DB |||||||
296 GAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGGCCACAGAGCCCAAGC 237
QY 5096 CTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTT 5155
DB |||||||
236 CTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTT 177
QY 5156 GCCTCTGGCTGTGTACAGGGTGAGCCCCAAAATTTGGGGTTTTCAGCGTGGGAGGCCACGTC 5215
DB |||||||
176 GCCTCTGGCTGTGTACAGGGTGAGCCCCAAAATTTGGGGTTTTCAGCGTGGGAGGCCACGTC 117
QY 5216 GATTTCTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGAGTAAAGTGAAGGA 5275
DB |||||||
116 GATTTCTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGAGTAAAGTGAAGGA 57

```

```

QY 5276 AGTTTATTAGAAAATAAAGAGGATATCATCAGCTCTTTTAGAA 5317
DB |||||||
56 AGTTTATTAGAAAATAAAGAGGATATCATCAGCTCTTTTAGAA 15

RESULT 15
LOCUS BM152462 655 bp mRNA linear EST 30-NOV-2001
DEFINITION TCBAP1E9080 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP9080, mRNA
sequence.
ACCESSION BM152462
VERSION BM152462.1 GI:17176729
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 655)
AUTHORS Wei,Y., Teang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
Gunaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
PUBLISHED
JOURNAL Unpublished
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Seq primer: M13 primer.
FEATURES
Location/Qualifiers
1. 655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP9080"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"
/note="vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTCAGCGCCGAGGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dc primer
[5'AGAGAGTCGGATCCGCGCGCAATAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 177 a 180 c 180 g 116 t 2 others
ORIGIN

Query Match 11.5%; Score 623.6; DB 12; Length 655;
Best Local Similarity 98.8%; Pred. No. 4.5e-74;
Matches 648; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1087 CCACCTCAGCCAGCTAGCACCAGAGAGGAGGCTCTGGACCCCAATGGCCTCTG 1146
DB |||||||
1 CCACCTCAGCCAGCTAGCACCAGAGAGGAGGCTCTGGACCCCAATGGCCTCTG 60
QY 1147 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAATCAGAG 1206
DB |||||||
61 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAAATCAGAG 120

```

QY	1207	AAAGGAGGCCCCCATGGGAGCGGTGGTAGGAACGCCCCCACAGGCGCACACCGCCTTA	1266
Db	121	ATTGGCAGGCCCCCATGGGAGCGGTGGTAGGAACGCCCCCACAGGCGCACACCGCCTTA	180
QY	1267	CAGCCCCACACACCCATGGGAGCGGTGGTAGGAGAGAGCGCTCTGTTCCACATGGGCC	1326
Db	181	CAGCCCCACACACCCATGGGAGCGGTGGTAGGAGAGAGCGCTCTGTTCCACATGGGCC	240
QY	1327	TGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTCTACAAAGACCTCAC	1386
Db	241	TGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTCTACAAAGACCTCAC	300
QY	1387	CCCAGAAGCCAAAGATCCCCTGGTCAAGAGAAGCTGGCCTGATTTATGTGGAGGAGATCGA	1446
Db	301	CCCAGAAGCCAAAGATCCCCTGGTCAAGAGAAGCTGGCCTGATTTATGTGGAGGAGATCGA	360
QY	1447	GGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCGCTGGATACCCAAAGACCTCGC	1506
Db	361	GGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCGCTGGATACCCAAAGACCTCGC	420
QY	1507	ATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCAGGCGAGGTGAAG	1566
Db	421	ATAGTCATACTGCAGGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCAGGCGAGGTGAAG	480
QY	1567	GAAGCCTGGGGGAGAGGCCAGCTGTATGGGGACCGCTTCCAGCATGTCTTCTACTTCAGC	1626
Db	481	GAAGCCTGGGGGAGAGGCCAGCTGTATGGGGACCGCTTCCAGCATGTCTTCTACTTCAGC	540
QY	1627	TGCAGAGAGCTGGCCCCAGTCCAAAGGTGGTGAAGTCTCGCTGAGCTCATCGGAAAAGATGGG	1686
Db	541	TGCAGAGAGCTGGCCCCAGTCCAAAGGTGGTGAAGTCTCGCTGAGCTCATCGG-AAAAGATGGG	599
QY	1687	ACAGCCACTCCGGCTCCCATTAGACAGATCCTGCTTAGGCCAGAGCGGCTGCTCTT	1742
Db	600	ACAGCCACTCCGGCT-CCATTAGACAGATCCTGCTTAGGCCAGAGCGGCTGCTCTT	654

Search completed: January 29, 2004, 13:05:49
Job time : 10855.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic. - nucleic search, using sw model

Run on: January 29, 2004, 03:02:31 ; Search time 291.391 Seconds
(without alignments)
8246.264 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggctggagagt.....aggataagaattacacac 5444

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.8	3.0	1371	2	US-08-910-731-1
2	165.8	3.0	1371	2	US-08-795-395-1
3	163.6	3.0	1371	2	US-08-910-731-7
4	163.6	3.0	1386	2	US-08-910-731-5
5	156.2	2.9	1374	2	US-08-910-731-3
6	156.2	2.9	1374	2	US-08-795-395-3
7	115	2.1	585	4	US-09-340-620A-50
8	115	2.1	740	4	US-09-340-620A-48
9	106.2	2.0	579	4	US-09-340-620A-62
10	106.2	2.0	777	4	US-09-340-620A-60
11	62	1.1	7218	1	US-08-232-463-14
12	59.4	1.1	257	4	US-09-016-434-208
13	57.2	1.1	3116	4	US-09-904-615-43
14	49	0.9	337	1	US-08-594-031-79
15	48.6	0.9	1878	4	US-09-465-558-39
16	48	0.9	1582	3	US-08-545-196B-10
17	48	0.9	1582	3	US-08-545-196B-12
18	47.2	0.9	17949	3	US-09-087-465-3
19	46.2	0.8	599	3	US-09-328-111-147
20	46.2	0.8	1441	3	US-08-821-994-63
21	46.2	0.8	4141	4	US-09-245-281-42
22	46.2	0.8	4141	4	US-09-207-359B-42
23	46.2	0.8	4141	4	US-09-340-620A-42
24	45.8	0.8	924	1	US-08-468-709B-1
25	45.8	0.8	924	2	US-08-241-664B-1
26	45.8	0.8	924	5	PCT-US93-03936-1
27	45.8	0.8	1717	1	US-08-468-709B-6

c 28	45.8	0.8	1717	2	US-08-241-664B-6	Sequence 6, Appli
c 29	45.8	0.8	1717	5	PCT-US93-03936-6	Sequence 6, Appli
c 30	45.6	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
c 31	45.2	0.8	289	3	US-09-007-005-17	Sequence 17, Appl
c 32	45.2	0.8	289	3	US-09-244-796-17	Sequence 17, Appl
c 33	45	0.8	1508	4	US-09-039-046-1	Sequence 1, Appli
c 34	45	0.8	11282	4	US-09-754-250-3	Sequence 3, Appli
c 35	44.4	0.8	1190	4	US-09-330-207-1	Sequence 1, Appli
c 36	44.4	0.8	3715	4	US-09-234-245-1	Sequence 1, Appli
c 37	43.8	0.8	467	2	US-08-841-349-18	Sequence 18, Appl
c 38	43.8	0.8	1868	1	US-08-658-883B-1	Sequence 1, Appli
c 39	43.8	0.8	1868	4	US-09-676-610B-26	Sequence 26, Appl
c 40	43.8	0.8	3437	3	US-08-860-339-17	Sequence 17, Appl
c 41	43.2	0.8	1342	4	US-09-489-847-89	Sequence 89, Appl
c 42	43.2	0.8	1534	1	US-08-300-903A-6	Sequence 6, Appli
c 43	43.2	0.8	1534	4	US-08-988-197-6	Sequence 6, Appli
c 44	42.8	0.8	1143	4	US-09-220-132-166	Sequence 166, App
c 45	42.8	0.8	1700	2	US-08-897-340-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-910-731-1
; Sequence 1, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHARTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS

i	LOCATION: 1...1368	Query Match	3.0%;	Score 165.8;	DB 2;	Length 1371;
US-08-910-731-1		Best Local Similarity	56.0%;	Pred. No. 8,7e-34;		
		Matches 314;	Conservative 0;	Mismatches 247;	Indels 0;	Gaps 0;
Qy	2920	CTCTTTCCCTCCTCAAGGTTCACACAGAAACCTGAAGGAGCTGGACCTTAATGTGAAACTCG	2979			
Db	811	CTCTGCGGTCTCCTCCAGGCGCAAGGAGACCTGAAGGAGCTCACTGTGGCGGGCAACAAG	870			
Qy	2980	CTGAGCCACTCTCAGTGAAGAGTCTTTGTAAAGACCTTGAGAGCCCTCGCTCGCTCCCTG	3039			
Db	871	CTGGGCGACAGGGCGCCCGGCTGCTGTGGAGAGCCTGTGTGAGCCCGGCTGCCAGCTG	930			
Qy	3040	GAGACCCCTGGCGTTGGCTGTGGGCTTCACAGCTGAGGACTGCAAGGACCTTGGCCTTT	3099			
Db	931	GAGTCCCTGTGGGTGAAGTCTTCAGCCTCAGGGGCGCTGCTGCCAGCAGCTCAGCTTG	990			
Qy	3100	GGGCTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGTTCATATGTCTACGGAT	3159			
Db	991	ATGCTGACCCAGAAACAAGCATCTCTGGAACTTCAGTTGAGCAGCAACAAGCTGGGCTGAC	1050			
Qy	3160	GCTGGAGCCAAAACACTTTGCCAGAGACTCAGACAGCCGAGCTGCAAGCTACAGCGACTG	3219			
Db	1051	TCGGCATCCAGAGCTGTGCCAGGCCCTGAGCAGCCGGGCAACACACTGCGGGTGCTC	1110			
Qy	3220	CAGCTGCTCAGCTGTGGCCCTCAGCTGTGCTGTGCCAGGACCTGGGCTCTGTGCTTAGT	3279			
Db	1111	TGTCCTGGGACCTGTGAGGTGACCAACAGCGGCTGACAGACCTCGCCTCGCTCTGCTG	1170			
Qy	3280	GCCAGCCCCAGCTGAAGGAGCTAGACCTGCAGCAGAAACCTGTGATGACGTTGGCGTG	3339			
Db	1171	GCCAAACGCGAGCTGCGAGAGCTGGACCTGAGCAACAACTGTGTGGCGCACCCGGGCGTC	1230			
Qy	3340	CGACTGCTCTGAGGGGCTCAGGCATCTCTGCTGCAAACTCATACGCTGGGGCTGGAC	3399			
Db	1231	CTCAGCTGTGTGGGAGCCTTGGAGCAGCCGGGCTGGCCCTGGAGCAGCTGGTCTCTGTATC	1290			
Qy	3400	CAGACAACTCTGAGTGATGATGAGTGAAGCAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCT	3459			
Db	1291	GACACCTACTGGACGGAGAGGTGGAGGACCGGCTTCAGGCCCTTGGNGGGGACCAAGCCC	1350			
Qy	3460	CAGCTGCTATCTTCAGCAGA	3480			
Db	1351	GGCCTGAGGCTCATCTCCTGA	1371			

RESULT 2
US-08-795-395-1
; Sequence 1, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997

	CLASSIFICATION:	435	
	PRIOR APPLICATION DATA:	60/024,057	
	FILING DATE:	16-AUG-1996	
	ATTORNEY/AGENT INFORMATION:		
	NAME:	ESMOND, ROBERT W.	
	REGISTRATION NUMBER:	32,893	
	REFERENCE/DOCKET NUMBER:	0942.3440002	
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE:	202-371-2600	
	TELEFAX:	202-371-2540	
	INFORMATION FOR SEQ ID NO:	1:	
	SEQUENCE CHARACTERISTICS:		
	LENGTH:	1371 base pairs	
	TYPE:	nucleic acid	
	STRANDEDNESS:	both	
	TOPOLOGY:	both	
	MOLECULE TYPE:	cDNA	
	FEATURE:		
	NAME/KEY:	CDS	
	LOCATION:	1..1368	
	US-08-795-395-1		
	Query Match	3.0%; Score 165.8; DB 2; Length 1371;	
	Best Local Similarity	56.0%; Pred.No.8.7e-34;	
	Matches 314; Conservative	0; Mismatches 247; Indels 0; Gaps 0;	
QY	2920	CTCTTTCGCTCCTCAAGGTCCACGAAACCCTGAAGGAGCTGGACCTAAGTGGAACCTCG	2979
DB	811	CTCTGCGGTGCTCTCAGGCCAAGGAGACCCCTGAAGGAGCTCAGTCTGGCGGGCAAACAAG	870
QY	2980	CTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCCTGAGACGCCCTCGCTGCCTCCCTG	3039
DB	871	CTGGGCGACGAGGCGCCCGGCTGCTGTGGAGAGCTCTGCAGCCCCGCTGCCAGCTG	930
QY	3040	GAGACCTCGGGTGTGCTGGCTGTGGCCCTACAGCTGAGGACTGCAAGAACCTTGGCTTT	3099
DB	931	GAGTCCCTGTGGGTGAAGTCTCTGCAGCCTCACGGCGGCTGTCTGCCAGCACGTCAGCTTG	990
QY	3100	GGGCTCAGAGCCAACAGACCCCTGACCGAGCTGGACCTCAGCTTCAATGTGTCACGGAT	3159
DB	991	ATGCTGACCCAGNACAAGCATCTCTGGAACTTCAGTTGAGCAGNACAAGCTTGGGTGAC	1050
QY	3160	GCTGGAGCCAAAACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGCACTG	3219
DB	1051	TCTGGCATCCAGAGCTGTGCCAGGCGCTTAGCCAGCCGGGCAACACACTGGCGGTGCTC	1110
QY	3220	CAGCTGGTCACTGTGGCCTCAGTCTGACTGTGTCAGGACCTGGCCTCTGTGCTTACT	3279
DB	1111	TGTCTTTGGGACTGTGAGGTGACCAACAGCGGCTGCAGCAGCTTCGCTTCCTGCTG	1170
QY	3280	GCCAGCCCAGCCTGAAGGAGCTAGACCTGCAGCAGAAACAACCTGGATACGCTTGGCGTG	3339
DB	1171	GCCAAACCGAGCCTGCAGAGCTGGACCTGAGCAACAACCTGTGGCGACCCGGGCGTC	1230
QY	3340	CGACTCTCTGTAGGGGCTCAGGCATCTCTGCTGCMAAATCATACGCTTGGGGCTGGAC	3399
DB	1231	CTGCAGCTCTGCGGGAGCCTGGAGCAGCGCGGCTGCGCCTCGAGCAGCTGCTCTGTAT	1290
QY	3400	CAGACAACCTGAGTCATCAGATGAGGAGGAACTGAGGGCCCTTGAGCAGGAGAACCT	3459
DB	1291	GACACCTACTGACGAGGAGGTGAGGAGACCGCCTGCAGGCCCTTGAGGGGAGCAAGCCC	1350
QY	3460	CAGCTGCTCATCTTCAGCAGA	3480
DB	1351	GGCCTGAGGGTCACTCTCTGA	1371

RESULT 3
US-08-910-731-7
; Sequence 7, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:

Query Match	3.0%;	Score 163.6;	DB 2;	Length 1371;
Best Local Similarity	55.7%;	Pred. No. 3.3e-33;		
Matches 313;	Conservative 0;	Mismatches 249;	Indels 0;	Gaps 0;
QY	2919	TCTCTTCTCGTCTCTCAAGTGCACCGAACCTGAAGGAGCTGGACCTAAGTGGAACTC	2978	
DB	810	TCTGTGCGGTCTCTCAGGCCAAGGAGAGCTCAAGGAGCTCAGCTTGGCCGGCAACGA	869	
QY	2979	GCTGAGCCACTCTCAGCTGAAGAGCTTTTGTAAACCCCTGAGACGCCCTCGCTGCCTCCT	3038	
DB	870	GCTGGGGATGAGGGTCCCGACTGCTGTGTGAGACCCCTGCTGGAACCTGGCTGCACGT	929	
QY	3039	GGAGACCCCTGCGTTTGGCTGGCTGTGSCCTCACAGCTGAGAGACTGCAAGGACCTTGCCTT	3098	
DB	930	GGAGTCGCTGTGGGTGAAGTCTCTGCAGCTTACAGACGCCCTGCTGCCCCCACTTCAGCTC	989	
QY	3099	TGGGCTGAGAGCCCAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGA	3158	
DB	990	AGTGCTGGCCCCAGAACAGGTTTCTCTGGAGCTACAGATAAGCAACACAGGCTTGGAGGA	1049	
QY	3159	TGCTGGAGCCAAACACCTTTTGCCAGACACTGAGACAGCCGAGCTGCAAGCTACAGCCACT	3218	
DB	1050	TGCGGGCTGGGGAGCTGTGCCAGGCGCTGGGCCAGCCTGGCTCTGTGCTGCGGGTGCT	1109	
QY	3219	GCAGCTGGTCAGCTGTGGCCTCAGTCTGACTGTGTGCAAGACTGTGGCCTCTGTGCTTAG	3278	
DB	1110	CTGGTTGGCCGACTGCCATGTGAGTGACAGCAGCTGCAAGAGCCTCGCCGCAACCCCTGTT	1169	

Qy 2919 TCTCTTCTCCGCTCCTCAAGGTCAACGAAACCTGAAGAGCTGGACCTAAGTGGAACTC 2978

```

Db      825  TCTGTCCCGTGTCTCAGGCGCCAAAGAGAGCCTGAAGAGAGCTCAGCCTGCGCGCAACGA 884
QY      2979  GCTGAGCCACTCTGCACTGAAGAGTCTTTGTAAGACCTGAGAGCGCCTCGTGTGCTCTCT 3038
Db      885  GCTGGGGGATGAGGTGCCGACTGTGTGAGACCTCTGCTGGAACCTGGCTGCAGCT 944
QY      3039  GGAGACCTCGCGTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3098
Db      945  GGAGTGGCTGGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
QY      3099  TGGGCTGAGAGCCAAACAGACCTGACCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 3158
Db      1005  AGTGTGGGCGCCAGAACAGAGTCTTCTCTGAGCTTACAGATAAGCAACAGAGCTGAGGA 1064
QY      3159  TGTGTGAGCCAAACACCTTTGCGAGAGCTGAGACAGCGAGCTGCAAGCTACAGGACT 3218
Db      1065  TGGGCGGTGGGAGCTGTGCGAGGCTGTGCGAGGCTGTGCGAGGCTGTGCGAGGCTGTGCGAGGCTGTG 1124
QY      3219  GCAGCTGGTCACTGTGGCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3278
Db      1125  CTGTTGGCGGACTGCGATGTGATGACAGAGCTGACAGAGCTGCGCGCAACCTGTT 1184
QY      3279  TGCAGCCCGAGCTGAAGAGCTAGACCTGAGAGAGCAACCTGAGAGCAACCTGAGAGCAACCTGAGAG 3338
Db      1185  GGCACACCAAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1244
QY      3339  GCGACTGCTGTGAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 3398
Db      1245  CTGCGAGCTGTGAGAGGCTGCGGAGCGGCTGCGGAGCGGCTGCGGAGCGGCTGCGGAGCGGCTGCG 1304
QY      3399  CCAGACAACTCTGAGTGATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3458
Db      1305  CGACATTTACTGCTGAGAGATGAGGAGCGGCTGAGGAGCGGCTGAGGAGCGGCTGAGGAGCGGCTGAG 1364
QY      3459  TCAGTGTCTATCTTCAGCAGA 3480
Db      1365  ATCCCTGAGGGTCACTCTCTGA 1386

```

```

RESULT 5
US-08-910-731-3
; Sequence 3, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1374 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
; US-08-910-731-3

Query Match      2.9%; Score 156.2; DB 2; Length 1374;
Best Local Similarity 54.2%; Pred. No. 3.1e-31;
Matches 317; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

QY      2896  GTCACAGATGCTATTGGCAGATTCTCTTCTCCGCTCTCAAGGTCACCCAGAAACCTGAAG 2955
Db      787  GTCACATGCAGAAAGGCTGCAAGGACCTGTGCCGTGTCTCTCAGAGCCAAGCAGAGCCTGAAG 846
QY      2956  GAGCTGGACCTTAAGTGGNAACCTCGCTGAGCCACTCTGCTGAGTGAAGAGCTTTTGTAAAGACC 3015
Db      847  GAACCTCAGCTAGCTGGCAATGAGCTGAAGATGAGGTGCCCAACTCTCTGTGTAGAGC 906
QY      3016  CTGAGACCCCTCGCTGCTCTCTGAGACCTCGCTGGGCTGTGGCTGTGGCTCTCACAGCT 3075
Db      907  CTGTAGAGCTGTGCTGTGCTGCTGAGTCACTGTGGGTAAAGACTGTAGCTCACAGCT 966
QY      3076  GAGGACTGCAAGGACCTTTGCTTTGGGCTGAGAGCAACACAGACCTTGACCGAGCTGAC 3135
Db      967  GCCTCTTGTCCCACTTCTGCTGCTGTGACCAAAACAGTTCTCTGTTGAGTTGCA 1026
QY      3136  CTGAGCTTCAATGTGCTCACGGATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAG 3195
Db      1027  ATGAGCAGCAACCCGCTGGGAGACTCGGAGCTTTTGAAGGCCCTGGGCTAT 1086
QY      3196  CCGAGCTGCAAGCTACAGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3255
Db      1087  CCGNACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
QY      3256  CAGGACCTGGCTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3315
Db      1147  AGCAGCCTTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
QY      3316  AACCACTGGATGAGCTTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3375
Db      1207  AACTGCATGGGGGCAACCGGTGCTCTCAACTGTGGAGGCTCAACAGCCAGCTGCTGCTGCTGCTGCTGCT 1266
QY      3376  AAACCTATACGCTGGGCTGGAGCCAGCAAACTCTGAGTGATGAGATGAGTGAAGCAAGCACTG 3435
Db      1267  ATCCTTACAGAGCTTGTCTGTATGACATTTTCTGAGCGGATGAGTGAAGCAAGCACTG 1326
QY      3436  AGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db      1327  CGGGCCCTGGAGGAGAAAGGCCATCCCTGAGGATCATTTCTCTGA 1371

```

```

RESULT 6
US-08-795-395-3
; Sequence 3, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor

```

; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/795,395
 ; FILING DATE: 04-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/024,057
 ; FILING DATE: 16-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ESMOND, ROBERT W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0942.3440002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1374 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1368
 ; US-08-795-395-3

Query Match 2.9%; Score 156.2; DB 2; Length 1374;
 Best Local Similarity 54.2%; Pred. No. 3.1e-31;
 Matches 317; Conservative 0; Mismatches 268; Indels 0; Gaps 0;
 QY 2896 GTCACAGATGCGCTATTGGCAGATTCTTCTCGCTCCTCAAGGTCAACAGACCTGGAAG 2955
 DB 787 GTCACTGCAGAGGCTGCAAGGACCTGTGCGGTGCTCTCAGAGCCAGCAGGCTGAG 846
 QY 2956 GAGCTGGACCTAAGTGGAACTCGCTGAGCCACTGTGACAGTGAAGAGTCTTTGTAAGACC 3015
 DB 847 GAACTCAGCTAGCTGGCAATGAGCTGAAGGATGAGGGTGCCCACTGCTGTGTGAGAGC 906
 QY 3016 CTGAGAGCGCCCTCGCTGCTCTGAGAGCCCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 3075
 DB 907 CTGTTAGAGCTGGCTGTGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 966
 QY 3076 GAGGACTGCAAGGACCTGCTGCTTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGAC 3135
 DB 967 GCCTCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
 QY 3136 CTGAGCTTCAATGTGCTACGAGATGTGAGGACCAACACTTTGTCAGAGACTGAGACAG 3195
 DB 1027 ATGAGCAGCAACCCGCTGGGAGACTCGGAGTCTGGAGCTTTGCAAGGCCCTGGGCTAT 1086
 QY 3196 CCGAGCTGCAAGCTAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3255
 DB 1087 CCGGACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
 QY 3256 CAGGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3315
 DB 1147 AGCAGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
 QY 3316 AACCACTGGATGACCTGGCGGTGCGAGTGTCTGTGAGGGGCTCAGGATCCTGCTGCTGCTGCT 3375

DB 1207 AACTGCATGGGGACAACGGTGTCTTCACTGCTGGAGAGCCTCAAAACAGCCAGCTGC 1266
 QY 3376 AAACCTATACGCTGGGCTGGACCCAGACAACCTCTGAGTGTGATGATGAGCAGGAACTG 3435
 DB 1267 ATCCTTCAGACGCTGTCTGTATGACATTTTACTGGACGATGAGTGTGAAGACCGACTT 1326
 QY 3436 AGGCGCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480
 DB 1327 CGGCGCTGGAGAGGAAGGCCATCCCTGAGGATCATTTCTCTGA 1371

RESULT 7

US-09-340-620A-50
 ; Sequence 50, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; CURRENT FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US 09/245,281
 ; PRIOR FILING DATE: 1999-02-05
 ; PRIOR APPLICATION NUMBER: US 09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 585
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-340-620A-50

Query Match 2.1%; Score 115; DB 4; Length 585;
 Best Local Similarity 61.5%; Pred. No. 1.5e-20;
 Matches 184; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
 QY 4476 CCCTCAGCCGCTAGCCGTACCTTCACTCTGATGCCCGCAGTGTGCTGCTGCTGCTGCT 4535
 DB 285 CGCGCAGCTGGATCCAGGCGCCCTCTCAGTCGACGAGCCAGCCAGGCTGCACTTTAT 344
 QY 4536 GGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGGTTGCTCTGCAAA 4595
 DB 345 AGACCAAGCAGCCGCTGCTGCTTATCGGAGGCTCAAAACGTTGAGTGGCTGCTGATGC 404
 QY 4596 ACTGCATGACAGAGTGTGAGCCAGGAGCAGTACGAGAGGCTGCTGGCTGAGAAACACGAG 4655
 DB 405 TCTGTACGGGAAGTCTTGTGCGATGAGCAGTACCCAGCAGTGCAGGCGCGAGCCCAAA 464
 QY 4656 GCCAGCAGATGCGGAAGCTGTTGAGTGTGAGCAGTCTGCGGAGCCGGAAGTGAAGA 4715
 DB 465 CCCAAGCAAGATGCGGAAGCTCTTCAGTTTCAACACGAGCTGGAACCTGACCTGCAAGA 524
 QY 4716 TGGACTCTACCAAGCCCTGAAGGAGACCCATCTCACCTCATTTATGGAACCTGCGAGA 4774
 DB 525 CTGTCTCTCCAGGCGCCCTAAGGAGTCCCACTTCTACTTGTGGAGGACCTGGAGCGGA 583

RESULT 8

US-09-340-620A-48
 ; Sequence 48, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; CURRENT FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US 09/245,281

;; PRIOR FILING DATE: 1999-02-05
;; PRIOR APPLICATION NUMBER: US 09/207,359
;; PRIOR FILING DATE: 1998-12-08
;; PRIOR APPLICATION NUMBER: US 09/099,041
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: US 09/019,942
;; PRIOR FILING DATE: 1998-02-06
;; NUMBER OF SEQ ID NOS: 71
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 48
;; LENGTH: 740
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (54)...(638)
US-09-340-620A-48

Query Match 2.1%; Score 115; DB 4; Length 740;
Best Local Similarity 60.9%; Pred. No. 1.8e-20;
Matches 187; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 4476 CCTCCAGCCCGCATACCGTACCTTCACCTCTGGATGCCCGCAGTTCCTGCACTTTGT 4535
DB 338 CGCCGAGCTGGGATCCAGGCCCTCTCAGTCGGCAGCCAAAGCCAGGCGCTGCACCTTAT 397
QY 4536 GGACCAAGTATCGAGCAGCTGATACCCGAGTGACATCGGTGGAGTTGTCTTGACAA 4595
DB 398 AGACCAAGCACCAGGCTGCGCTTATCCGAGGGTTCACAAACGTTGAGTGTCTGATGC 457
QY 4596 ACTGCATGGACAGGTCTGTCAGCCAGCAGTACGAGAGGGTCTGCCTTGAGAACACAG 4655
DB 458 TCTGTACGGGAAGTCTCTACGGATGAGCAGTACGAGGAGTGGCGGCGAGCCACCAA 517
QY 4656 GCCAGCCAGATCGGAAGCTGTTACGTTGAGCCAGTCTCGGACCGGAAGTGCAAGA 4715
DB 518 CCCAAGCAAGATCGGAAGCTCTCAGTTTCACACCCAGGCTGGAACCTGGAAGGA 577
QY 4716 TGGACTCTACCAAGCCCTGAGGAGACCCATCTCACCCTCATTTATGGAACCTGGAGAA 4775
DB 578 CTTGCTCTCCAGGCCCTTAAAGGAGTCCCAAGTCTTACCTGGTGGAGGACCTGGAGCGGAG 637
QY 4776 GGGCAGC 4782
DB 638 CTGAGGC 644

RESULT 9
US-09-340-620A-62
;; Sequence 62, Application US/09340620A
;; Patent No. 6482933
;; GENERAL INFORMATION:
;; APPLICANT: Bertin, John
;; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;; FILE REFERENCE: 07334-124001
;; CURRENT APPLICATION NUMBER: US 09/340,620A
;; CURRENT FILING DATE: 1999-06-28
;; PRIOR APPLICATION NUMBER: US 09/245,281
;; PRIOR FILING DATE: 1999-02-05
;; PRIOR APPLICATION NUMBER: US 09/207,359
;; PRIOR FILING DATE: 1998-12-08
;; PRIOR APPLICATION NUMBER: US 09/099,041
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: US 09/019,942
;; PRIOR FILING DATE: 1998-02-06
;; NUMBER OF SEQ ID NOS: 71
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 62
;; LENGTH: 579
;; TYPE: DNA
;; ORGANISM: Mus musculus
US-09-340-620A-62

Query Match 2.0%; Score 106.2; DB 4; Length 579;
Best Local Similarity 64.4%; Pred. No. 3.3e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 4528 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 4587
DB 331 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 390
QY 4588 TTGGACAAACTGTCATGGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTGCTGCTGAG 4647
DB 391 CTGGATGCTTTGCAATGGCAGTGTGCTGACTGAAGACAGTACCGGCAAGTTCGTGCAGAG 450
QY 4648 AACACGAGCCCGCAGCCAGATCGGAAGCTGTTTCAGCTTTGAGCCAGTCTCTGGACCGGAAG 4707
DB 451 ACCACCAAGCAAGCAAGATGAGGAAGCTCTTCAGCTTTGTTCCATCTCTGGAACCTGACC 510
QY 4708 TGCAAAGATGGAAGTCTTACCAAGCCCTGGAAGAGACCCATCTCTACCTCATTTATGGAATC 4767
DB 511 TGCRAAGGACTCCCTCTCCAGGCCCTTGAAGGAAATACATCCCTACTTGTGATGAGACCTG 570
QY 4768 TGGGAGA 4774
DB 571 GAGCAGA 577

RESULT 10
US-09-340-620A-60
;; Sequence 60, Application US/09340620A
;; Patent No. 6482933
;; GENERAL INFORMATION:
;; APPLICANT: Bertin, John
;; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;; FILE REFERENCE: 07334-124001
;; CURRENT APPLICATION NUMBER: US 09/340,620A
;; CURRENT FILING DATE: 1999-06-28
;; PRIOR APPLICATION NUMBER: US 09/245,281
;; PRIOR FILING DATE: 1999-02-05
;; PRIOR APPLICATION NUMBER: US 09/207,359
;; PRIOR FILING DATE: 1998-12-08
;; PRIOR APPLICATION NUMBER: US 09/099,041
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: US 09/019,942
;; PRIOR FILING DATE: 1998-02-06
;; NUMBER OF SEQ ID NOS: 71
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 60
;; LENGTH: 777
;; TYPE: DNA
;; ORGANISM: Mus musculus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (89)...(667)
US-09-340-620A-60

Query Match 2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 4e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 4528 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 4587
DB 419 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 478
QY 4588 TTGGACAAACTGTCATGGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTGCTGCTGAG 4647
DB 479 CTGGATGCTTTGCAATGGCAGTGTGCTGACTGAAGACAGTACCGGCAAGTTCGTGCAGAG 538
QY 4648 AACACGAGCCCGCAGCCAGATCGGAAGCTGTTTCAGCTTTGAGCCAGTCTCTGGACCGGAAG 4707
DB 539 ACCACCAAGCAAGCAAGATGAGGAAGCTCTTCAGCTTTGTTCCATCTCTGGAACCTGACC 598
QY 4708 TGCAAAGATGGAAGTCTTACCAAGCCCTGGAAGAGACCCATCTCTACCTCATTTATGGAATC 4767
DB 599 TGCRAAGGACTCCCTCTCCAGGCCCTTGAAGGAAATACATCCCTACTTGTGATGAGACCTG 658

[illegible][illegible]

Db 62 GAGACGAGACGCTGTTGCGCTGTGCTTGTGCTCCCTTGTGCTGCTGATCGTGTGCACC 121
QY 2107 TGCGCTGATCAGCAGATGAGCGGAGGAGAAACTCACACTGACTTCCAGACCCACCA 2166
Db 122 GTGCTG-CGCGANANTGNGAGCTCGGTGCGGACCTGTGCGCACGTCCAGACCCACG 180
QY 2167 ACCTCTGTCTACATTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2198
Db 181 TCAGTGTACTGCTTTTTCATCACCAGCGTCT 212

RESULT 13

US-09-904-615-43
; Sequence 43, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 43
; LENGTH: 3116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-615-43

Query Match 1.1%; Score 57.2; DB 4; Length 3116;
Best Local Similarity 50.9%; Pred. No. 9.4e-05;
Matches 162; Conservative 0; Mismatches 153; Indels 3; Gaps 1;

QY 3023 GCCTCGCTGCTCTGAGACCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082
Db 1036 GCACCTGCCACCACTTGGAGGAGCTGGACTTGTCTAACCAATCAATTTGATGAGGAGGCA 1095
QY 3083 GCAAGACCTTGCTTTGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTGAGCT 3142
Db 1096 CNAAGCGCTGATGAGGGCCCTTGAGGGGAAATGGATGTAAGAGGCTGGACCTCAGTC 1155
QY 3143 TCAATGTCTACGGATGTGGAGCCAAACACCTTTGCCAGAGACTGACACAGCCGAGCT 3202
Db 1156 ACCTTCTGTGACAGCTCCACCTTGGCTTGTCTTACTCACAGACTAAGCCAGATGACCT 1215
QY 3203 GCNAGCTACGCGACTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3262
Db 1216 GC---CTGACAGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
QY 3263 TGGCTCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3322
Db 1273 TTTCTGAGCTTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
QY 3323 TGGATGACGTTGGCGTGC 3340
Db 1333 TTGGAGACGCTGGTGTC 1350

RESULT 14

US-08-594-031-79
; Sequence 79, Application US/08594031
; Patent No. 578182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-594-031-79

Query Match 0.9%; Score 49; DB 1; Length 337;
Best Local Similarity 62.8%; Pred. No. 0.0035;
Matches 76; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 5317 ATTTCTGTAGCAGGCTTCCAGTCTTTTACAGAGAAACCCCTATATAATTAATAATTTT 5376
Db 215 ATTTATGATATGTTTGTCTACTCTCTGACACATGCTTATAACTGAATTAGGCAGT 274
QY 5377 CTTAAATTTAAGAAATTAATAATAATACAAAAAGAAAAATGAAATAAAGGAATGAAG 5436
Db 275 CTTAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAC 334
QY 5437 T 5437
Db 335 T 335

RESULT 15

US-09-465-558-39
; Sequence 39, Application US/09465558
; Patent No. 6436657
; GENERAL INFORMATION:
; APPLICANT: Morakinyo, Lavo O.
; APPLICANT: Orozco Jr, Emil M.
; TITLE OF INVENTION: TETRAHYDROFOLATE METABOLIC ENZYMES
; FILE REFERENCE: BBI322 US NA
; CURRENT APPLICATION NUMBER: US/09/465,558
; CURRENT FILING DATE: 1999-12-17
; EARLIER APPLICATION NUMBER: 60/112,734
; EARLIER FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 39
; LENGTH: 1878

```

; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-465-558-39

Query Match      0.9%; Score 48.6; DB 4; Length 1878;
Best Local Similarity 52.8%; Pred. No. 0.013;
Matches 105; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 5237 AGATCTACAAGAGCAAGCCACACAGAGTAAAGTGGAAAGGAGTTTATTTCAGAAAAATAAAGG 5296
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1672 AGATCCATGATGAGCGCCACCATATGTAAAGGATCCAGCATTTTACAGATATGGG 1731

QY 5297 AGTATCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTCCAGTTTTCACGAAACCC 5356
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1732 AACTTTGTCGATAGTTTCTTATTGCGAGGCACATACTGTAAGATGCTTCGCTGATATGCTA 1791

QY 5357 TATAAATTAAAAATTTTTTACTTTAAATTTAAGAAATTAAGAAATACAAAAAGAAAAAT 5416
    || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1792 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 1851

QY 5417 GAAATATAAGGAATAAGAA 5435
    |||| | | | | | | | |
Db 1852 AAAAAAATAAAAAAATAAAAAA 1870
```

Search completed: January 29, 2004, 13:11:27
Job time : 293.391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 35.197 Seconds
(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWGLACYLEFLKKE.....HLIMELWEKSKKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5969	79.2	1192	T17255	hypothetical prote
2	882	11.7	1111	A53000	matr protein (imp
3	465.5	6.2	483	S27880	Nasopressin recept
4	361.5	4.8	456	S20597	ribonuclease inhib
5	358	4.8	461	A31858	ribonuclease-angio
6	354.5	4.7	456	A31857	ribonuclease inhib
7	345	4.6	1130	A48943	MHC class II trans
8	206	2.7	1075	T31668	hypothetical prote
9	202	2.7	1004	T31665	hypothetical prote
10	198	2.6	1232	A55478	neuronal apoptosis
11	196.5	2.6	1447	T42628	neuronal apoptosis
12	180	2.4	533	T52063	ran GTPase-activat
13	166.5	2.2	1120	JC7765	mitotic spindle as
14	159.5	2.1	506	A45841	T-complex-associat
15	155.5	2.1	1199	T47442	disease resistance
16	153.5	2.0	1131	F96662	hypothetical prote
17	152	2.0	1027	T46296	hypothetical prote
18	152	2.0	3511	A59295	unconventional myo
19	151	2.0	1221	T52347	disease resistance
20	150.5	2.0	1220	T48928	disease resistance
21	149	2.0	631	C89243	protein F28C1.3 [i
22	149	2.0	631	T21471	hypothetical prote
23	148	2.0	1217	T52348	disease resistance
24	147.5	2.0	793	S54772	mammary gland fact
25	147.5	2.0	1214	T47438	disease resistance
26	147	2.0	1017	H96663	hypothetical prote
27	147	2.0	1784	T10532	sag-pol polyprotei
28	143.5	1.9	4302	A38971	polycystatic kidney
29	142	1.9	1215	T00364	hypothetical prote

30	142	1.9	1839	1	OYBYK	adenylate cyclase
31	139.5	1.9	545	2	T52068	RAN GTPase-activat
32	139.5	1.9	1189	2	T52346	disease resistance
33	138.5	1.8	1031	2	G96663	hypothetical prote
34	138	1.8	1010	2	T36383	probable large ATP
35	137	1.8	1226	2	T45788	probable disease r
36	136.5	1.8	794	2	G23317	transcription acti
37	136.5	1.8	3335	2	H81702	adherence factor 1
38	135.5	1.8	1422	2	B71437	probable resistanc
39	135	1.8	915	2	T09575	smoothelin - human
40	135	1.8	1724	2	T18343	p-glycoprotein - S
41	134	1.8	889	2	H96606	hypothetical prote
42	133	1.8	1184	2	H71436	hypothetical prote
43	133	1.8	1301	2	D85188	disease resistance
44	133	1.8	1453	2	G96613	hypothetical prote
45	132.5	1.8	805	2	AH2731	ATP-dependent prot

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFp586O1822.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17255
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: EMBL:AL117470
A:Experimental source: adult uterus; clone DKFp586O1822
C:Genetics:
A:Note: DKFp586O1822.1

Query Match		79.2%; Score 5969; DB 2; Length 1192;	
Best Local Similarity		96.2%; Pred. No. 0;	
Matches 1147; Conservative 1; Mismatches 0; Indels 44; Gaps 1;			
QY	282	LLQLRPHRSDPLVKRSDPYVEENRGLHIEIRDLFGPLDTQEPRIVLQGAAGIGKS	341
DB	1	LLQLRPHRSDPLVKRSDPYVEENRGLHIEIRDLFGPLDTQEPRIVLQGAAGIGKS	60
QY	342	TLARQVKEAWRGQLYGDRFQHVFFSCRELAQSKVSLAEHLIGKDGATATAPIRQILSR	401
DB	61	TLARQVKEAWRGQLYGDRFQHVFFSCRELAQSKVSLAEHLIGKDGATATAPIRQILSR	120
QY	402	PERLLFILDGVDEPGWVLOEPSSELCLHWSQOPADALLGSLGKLTILPEASFLITARTT	461
DB	121	PERLLFILDGVDEPGWVLOEPSSELCLHWSQOPADALLGSLGKLTILPEASFLITARTT	180
QY	462	ALQNLIPSLQEARWVEVLGFSSESSKEYFYFYFTDERQAIRAFRLVKSNKELWALCLVPW	521
DB	181	ALQNLIPSLQEARWVEVLGFSSESSKEYFYFYFTDERQAIRAFRLVKSNKELWALCLVPW	240
QY	522	VSWLACTCLMQMKRKEKLTLSKTTTTLCHLYLAQALQAPGLQPLDCLSLAAGIQQ	581
DB	241	VSWLACTCLMQMKRKEKLTLSKTTTTLCHLYLAQALQAPGLQPLDCLSLAAGIQQ	300
QY	582	KTTLFSPDRLRKHGDGALISTFLKMGILQEHPIPLSYFHLCHFQEPFAAMSYYLEDEK	641
DB	301	KTTLFSPDRLRKHGDGALISTFLKMGILQEHPIPLSYFHLCHFQEPFAAMSYYLEDEK	360
QY	642	GRGKHSNCIIDLKTEKLEAYGHLFGCASTRFLGLLSDEGEREMENIFHCRLSQGRNLM	701
DB	361	GRGKHSNCIIDLKTEKLEAYGHLFGCASTRFLGLLSDEGEREMENIFHCRLSQGRNLM	420
QY	702	QWVPSLQLLQPHSLLESCLYETRNKFTLTQVMAHFEMGMCVETDMELLVCTFCIKPS	761

Db 421 QWVPSLQLLQPHSLBSLHCLYETRNKTLFTQVMAHFEEMGCMVETDMELLVCTTCIKFS 480
 Qy 762 RVVKLQLLEGQRHRTWSPMTVVLFRWVPVTDAYWQIILFVSLKVRNKLKELDLGNSLS 821
 Db 481 RVVKLQLLEGQRHRTWSPMTVVLFRWVPVTDAYWQIILFVSLKVRNKLKELDLGNSLS 540
 Qy 822 HSAVSKLCTLRPRCLLETTLRAGCGLTAEDCKOLAFGLRANQTLTLDLSFNVLTDAG 881
 Db 541 HSAVSKLCTLRPRCLLETTLRAGCGLTAEDCKOLAFGLRANQTLTLDLSFNVLTDAG 600
 Qy 882 AXHLQRLQPSCKLQRLQVSCGLTSDCCDLASVLSASPSLKLKELDLQNNLDDVGVRL 941
 Db 601 AXHLQRLQPSCKLQRLQVSCGLTSDCCDLASVLSASPSLKLKELDLQNNLDDVGVRL 660
 Qy 942 LCEGLRHPACKLIRGLDQTTLSDEMRQBLRALEOEKPOLLIIFSRKPSVMTPTGELDTG 1001
 Db 661 LCEGLRHPACKLIRGLDQTTLSDEMRQBLRALEOEKPOLLIIFSRKPSVMTPTGELDTG 720
 Qy 1002 EMSNSTSSLRQRGLSERAASHVAQANLKLDDVSKIIFPIAEIAESSPEWVVELLCVPS 1061
 Db 721 EMSNSTSSLRQRGLSERAASHVAQANLKLDDVSKIIFPIAEIAESSPEWVVELLCVPS 780
 Qy 1062 PASQGLHTKPLGTDGDFGPTGPVATEVVDKKNLYRVHFPVAGSYRWNTGLCFVWRE 1121
 Db 781 PASQGLHTKPLGTDGDFGPTGPVATEVVDKKNLYRVHFPVAGSYRWNTGLCFVWRE 840
 Qy 1122 AVTVEIEFCVWDFLGEINPOHSMWVAGPLDIDKAPGAVEAVHLPHFVALOGGHVDTSL 1181
 Db 841 AVTVEIEFCVWDFLGEINPOHSMWVAGPLDIDKAPGAVEAVHLPHFVALOGGHVDTSL 900
 Qy 1182 FQVAHKEGMLLEKPARVELHIVLENPFSPLGVLLKMIHNAIRFIIVTSVLLYHRV 1241
 Db 901 FQVAHKEGMLLEKPARVELHIVLENPFSPLGVLLKMIHNAIRFIIVTSVLLYHRV 960
 Qy 1242 HPEVTHLYLIPSDCSIR----- 1260
 Db 961 HPEVTHLYLIPSDCSIRKADIDLEMPQFVRIHKPPPLTPLYMGRYTVSGSGGML 1020
 Qy 1261 ---KELECYRSPGEDOLSEFVGHGSGIRLOVKDKDETLVWALVKPGDLMPATTL 1317
 Db 1021 ILPKELECYRSPGEDOLSEFVGHGSGIRLOVKDKDETLVWALVKPGDLMPATTL 1080
 Qy 1318 IPPARIVAPSPIDAPOLLHFVQYRQRLIARTVSEVWLDKLHGQVLSQEQYERVLAE 1377
 Db 1081 IPPARIVAPSPIDAPOLLHFVQYRQRLIARTVSEVWLDKLHGQVLSQEQYERVLAE 1140
 Qy 1378 RPSQMRKLFSLQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
 Db 1141 RPSQMRKLFSLQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1192

RESULT 2
 A59000
 mater protein [imported] - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: A59000
 R:Tong, Z.B.; Nelson, L.M.
 Endocrinology 140, 3720-3726, 1999
 A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ovulation
 A:Reference number: A59000; MUID:99360614; PMID:10433232
 A:Accession: A59000
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1111 <KUR>
 A:Cross-references: GB:AF074018; NID:G5802697; PIDN:AAD51762.1; PID:G5802698
 C:Genetics:
 A:Gene: Mater

Query Match 11.7%; Score 882; DB 2; Length 1111;
 Best Local Similarity 26.1%; Pred. No. 2.4e-49;
 Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

Qy 162 SSPDHESQSPSNAP-----TSTAVLGSWGGPPQPS-----LAPR-----EQAPGTQWPLD 209
 Db 28 TSPENDSKIQKDGPEQEQTSESTMG---PPEKDSKAILKARGLEBEQKSESTMSPE 83
 Qy 210 ETSGIYYTEIREREREKSEKGRPPWAAVVGTPPQATSLQPHHPHWPESVRESLSTWFW 269
 Db 84 NVSRAILKDSGSEVEQASERK-----MTSPENDSKSIQKDGPEQEQTSETLOS----- 133
 Qy 270 KNEDFNQFTQLLLQRLPHRPSQDPLVKRSPDVPVEENRGLIE----- 313
 Db 134 KEED---EVTE-----ADKNGGD-LQDYKAHVIAKFTDTSVDLHYDSEPMK 175
 Qy 314 -IRDLFGGLDQTBEPRIVILQAGIGKSTTLARQVKEAWRGOLYGDGFQHFVFSCHREL 372
 Db 176 LLSDAFKYQKTFQPHHTIILHGRPGVGSALARSIVLQWAGKLF-QKMSVFIPIESVREI 234
 Qy 373 AQSKVVSIAELIGKDGATAPAPIQLSRPRLILFDGVDEPGWVLQEPSELCLHWSQ 432
 Db 235 KWTEKSSILAQLTAKCEPDSDVLVTKIMSQPBRLLFVIDGLDDMDSVLQHDMDTLSDWKD 294
 Qy 433 POPADALLGSLGKLTILPEASFLITARTALONLIPSLQEARWVVEVLGFSSESSKEYFYR 492
 Db 295 EQPIYILMYSLLRKALLPQSFIIITRNTGLEKLSMWVSPILYILVEGLSASRRSQVLVE 354
 Qy 493 YFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQQMKRKEKLTLSKTTT----- 548
 Db 355 NISNESDRIQVPHSLIENHQLFDQCAQPSVSLVCEALQLOKLGKRCRCTLPCQTLTGLYA 414
 Qy 549 TILCHYL-----AQAQAQPLGQPLRDLCSLAAGIOWKKTLPSPDDLRKHGLDGAIST 603
 Db 415 TLVPHQLTKRPSQSALSQEBQITLVGLCMAAGVMTMRVSFYDDDLKNSLKESETLA 474
 Qy 604 FLKWLQLOE--HPILPSYSFTHLCFOBFPAAMSYYLEDEKGRKH-----SNCIIDLEK 655
 Db 475 LFHMNILLQVHNSEQCVFVSHLSLQDFFAALYYVLEGLEEWNQHFCEIENQSRIMEYKR 534
 Qy 656 TLEAYGIHGLFGASTRELLGLLSDGEREMENIFHCLR--SQRNLMQWPSLQLLLQ 713
 Db 535 TDDT---RLJGK--RFLFGLMKNKILKTLEVLFEYFVPIVEQKQHWV---SLIAQQ 585
 Qy 714 -----HSLBSLHCLYETRNKTLFTQVMAHFEEMGCMVETDMELLVCTTCIKSRHVK 766
 Db 586 VNGTSPMDTLDAFYCLFESQDEEFVGGALKRFQEWLLINQKMDLKVSYSYCLKHQNLKA 645
 Qy 767 LQL-----TEGQRHSTWSPMTVVL-PRWVPVTDAYWQIILFVSLKVRNKLKELDLGNSLS 821
 Db 646 IRVDIRDLSDVNTLELCPVVVTOETQCKPLLMWGNFCSVLGSLRNLKELDLGDSILS 705
 Qy 822 HSAVSKL----- 828
 Db 706 QRAMKILCLELRNQSRIQKLTFFKSAEVVSGLKLWKLFLSNQNLKYLNLGNTPMKDDDM 765
 Qy 829 ---CKTLRRPRCLLETLR----- 843
 Db 766 KLACEALKHPKCSVETLRLDSCELTIIGYEMISTILLITRCLKLSLAKNRVGVKSMISL 825
 Qy 844 -----LACGGLTAEDCKOLAFGLRANQTLTLDLSFNVLTDAGAKHLQCR 888
 Db 826 GNALSSMCLLQKLTLDNCGUTPASCHLLVSALFSNQNLTHLCLSNNSLGTGEGVQOLCOF 885
 Qy 889 LRQPSCKLQR----- 898
 Db 886 LNFECALQRLILNHCNIVDDAYGFLAMLANNTKLTLSLTMTNVPVGDAMKLLCEALKE 945
 Qy 899 -----LQVSCGLTSDCCDLASVLSASPSLKLKELDLQNNLDDVGVRLLCCEGLRHPAC 951
 Db 946 PTCVQLQELVDCQLTONCCEDLACMITTKHLKSLDLGNALGDKGVITTLCEGLKQSSS 1005
 Qy 952 KLIRGLDQTTLSDEMRQBLRALEOEKPOL 981
 Db 1006 SLRRLGLGACKLTSNCCCEALSALSCNPHL 1035

RESULT 3

S27880
Nasopressin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C:Accession: S27880
R:Herrera, V.L.; Ruiz-Opazo, N.
submitted to the EMBL Data Library, February 1992
A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
A:Reference number: S27880
A:Accession: S27880
A:Molecule type: mRNA
A:Residues: 1-483 <R>
C:Cross-references: EMBL:M85183; NID:g202805; PIDN:AAA03623.1; PID:g202806
C:Genetics:
A:Gene: AVP

Query Match 6.2%; Score 465.5; DB 2; Length 483;
Best Local Similarity 30.6%; Pred. No. 1.3e-22;
Matches 161; Conservative 74; Mismatches 186; Indels 105; Gaps 18;
QY 534 MKRKEKLTSTKTTTCLHLYAALQ-AQPLQP-----QLRDLCSLAAGIWKTKTLPSP 588
DB 1 MELGRDLSTSTKTTTSTVYLLFTSMKSAAGTPRVQGEMLCLRLAREGILKHQAFSE 60
QY 589 DDLR--KHGLDGAIIIST-FLKM---GILQEHPIPLSYFIHLCPQFFPAAMSYVLEDEK 641
DB 61 KDLERLKLKLGSGVQVQTMFLSKKELPGVLE---TVVTYQFIDQSFOEFLAALSYLLDAB- 116
QY 642 GRGKSNCIIDLEKTLKLEAGIHLGFLG--ASTFRFLGLLSDGEREMENIFHC-----RLS 695
DB 117 --GAPGNSAGSVQMLNSDA--GLRGHLALTTRFLFGLLSTERIDGNHFGCVWVRVK 172
QY 696 QGRNLQMWPSLQLLQP-----HSLSEHLCLY 723
DB 173 Q--DTLRWV---QQQPKVATVGAEBKDELKDEABEEEEEEEEELNFGLELLYLY 227
QY 724 ETRNKTFLQVMAHFEMGCVB---TDMELLVCTFCIKFSRHVKQLQI-----E 771
DB 228 ETQEDDFVQALSSLPB--WVLERVLTMDLEVLVSYVCCPDGQALRLVSCGLVAKE 285
QY 772 GQHRSTWPTWVLFPRWPTVDAYWQIILFVLKTRNLKELDLGNSLSHSAVSKLCT 831
DB 286 KKKKKKSF-----NNRLKGSQSTGKPPASILLRPLCEA 318
QY 832 LRPRCLLETLRAGCLTAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLQ 891
DB 319 MITQQCGSLTILSHCKLPDAVCRDSEALKVAPSLRELGLOQRNTEAGRLLSOGLAW 378
QY 892 PSCKLQRLQVSCGLTSDCCQDLASVLSASPSELKELDLQNNLDDVGVRLLCGLRHAPAC 951
DB 379 PKCKVQTLRIQMPGL-QEVTHYLVIVLQSPVLTLLDSCQQLPGTVVPELCSALKHPKC 437
QY 952 KLIRLGLDQTLSDENRQELRALEQBKPOLLIPIRRKPSVMTPEG 997
DB 438 GLKTLSTLSTENPLRELAQVKTLPDLAIHSLKLGTHPQLKG 483

RESULT 4

S20597
ribonuclease inhibitor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S20597
R:Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992
A:Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
A:Reference number: S20597; MUID:92162755; PMID:1536887
A:Accession: S20597
A:Molecule type: mRNA
A:Residues: 1-456 <RAW>
C:Cross-references: EMBL:X62528; NID:g57670; PIDN:CAA44388.1; PID:g57671
C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 4.8%; Score 361.5; DB 2; Length 456;
Best Local Similarity 35.5%; Pred. No. 7.9e-16;
Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;
QY 792 VTDAWQIILFVLKTRNLKELDLGNSLSHSAVSKLCTLRPRCLLETLRAGCLTA 851
DB 206 ITSANCKLDCDVVASKLQELDLGNSLGNLTGIAALCSGLLLPSCRLRTLWLPDCDVTA 265
QY 852 EDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLRQPSCKLQRLQVSCGLTS--- 908
DB 266 EGCKDLCRVLRKQSKLSLGNELKDEGAQLLCESLLEPCQQLSLWVKTCSLTAASC 325
QY 909 -----DC-----CQDL 914
DB 326 PHFCSVLTKNSLPFELQMSNPLGDSGVVELCKALGPDTVLRVLWLGCDVTDSCSSL 385
QY 915 ASVLSASPSLKELDLQNNLDDVGVRLLCGLRHAPACKLIRLGLDQTLSDENRQELRAL 974
DB 386 ATVLNARSLRELDLSNNCGDNGVLQLESKQPSCLQQLVLYDIYWTDEVEDQALAL 445
QY 975 EOEKPOLLIIS 985
DB 446 EERPSLRIIS 456

RESULT 5

A31858
ribonuclease-angiogenin inhibitor - human
N:Alternate names: ribonuclease inhibitor, placental
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000
C:Accession: A31858; S02012; S23933; S48636; T47188
R:Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
Biochemistry 27, 8545-8553, 1988
A:Title: Primary structure of human placental ribonuclease inhibitor.
A:Reference number: A31858; MUID:89118269; PMID:3219362
A:Accession: A31858
A:Molecule type: mRNA
A:Residues: 1-461 <LES>
A:Cross-references: GB:M22414; NID:g186260; PIDN:AA59130.1; PID:g307040
R:Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
EMBO J. 7, 4151-4156, 1988
A:Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) disclosed
A:Reference number: S02012; MUID:89210799; PMID:3243277
A:Accession: S02012
A:Molecule type: mRNA
A:Residues: 1-422, 'SE', 425-461 <SCH>
A:Cross-references: EMBL:X13973; NID:g35843; PIDN:CAA32151.1; PID:g35844
A:Note: part of this sequence, including the carboxyl end of the mature protein, was con
R:Crevel-Thieffry, I.; Cotterill, S.; Schuller, E.
Biochim. Biophys. Acta 1122, 107-112, 1992
A:Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibiti
A:Reference number: S23933; MUID:92338217; PMID:1633192
A:Accession: S23933
A:Molecule type: protein
A:Residues: 174-195 <CRE>
R:Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.
Arch. Biochem. Biophys. 312, 421-428, 1994
A:Title: Purification and characterization of human brain ribonuclease inhibitor.
A:Reference number: S48636; MUID:94311593; PMID:8037455
A:Accession: S48636
A>Status: preliminary
A:Molecule type: protein
A:Residues: 2-14 <RAD>
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24374
A:Accession: T47188
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <AAA>

A;Cross-references: EMBL:AL161967
A;Experimental source: adult testis; clone DKFp243K249

C;Genetics:

A;Gene: GDB:RNH

A;Cross-references: GDB:125274; OMIM:173320

A;Map position: lip15.5-1lip15.5

A;Note: DKFp243K249.1

C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

F;34-433/Region: leucine-rich 57-residue repeats

F;57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F;174-195/Region: inhibitory

F;228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F;399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Query Match 4.8%; Score 358; DB 2; Length 461;

Best Local Similarity 37.9%; Pred. No. 1.4e-15;

Matches 86; Conservative 30; Mismatches 91; Indels 20; Gaps 3;

QY 768 QLIEGQHRSWPTWVLFVFWVPTDAYWQI-----LFSVLKVTNKLKELDSLGN 818

DB 246 ELCPGLHPSSRLTL-----WT-----WECITAKGCDLCRVLRKESLKEISLAGN 294

QY 819 SLHSVAVKSLCKTLRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTTLELDLSFNVL 878

DB 295 ELGDEGARLLCETLLPFGQLESWLKWSCTFTAACCSHFSSVLAQNRFLLEIQISNRLE 354

QY 879 DAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSAPSKELDLQNNLDDVG 938

DB 355 DAGVRELCCGLGPGSVLRVLMADCDVSDSSCSSLAATLLANHSURELDLSNCLGDAG 414

QY 939 VRLLCGLRHPACKLRGLDQTLTSDENRQELRALEQKPOLLI 985

DB 415 ILQVESVRPGCLLEQLVLYDIWSEEMEDRLQALEKDKPSLRVIS 461

RESULT 6

A31857 ribonuclease inhibitor, hepatic - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000

C;Accession: A31857; A35830

R;Hofereenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.

Biochemistry 27, 8537-8544, 1988

A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals th

A;Reference number: A31857; MUID:89118268; PMID:3219361

A;Accession: A31857

A;Molecule type: protein

A;Residues: 1-456 <HOF>

R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;

Biochemistry 29, 8827-8834, 1990

A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea

A;Reference number: A35830; MUID:91104783; PMID:2271559

A;Accession: A35830

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 82-456 <VIC>

A;Cross-references: GB:M58700; NID:G164638; PIDN:AAAG3448.1; PID:G164639; GB:J02925

C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

C;Keywords: liver

F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match

Best Local Similarity 34.3%; Pred. No. 2.3e-15;

Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

QY 752 LVCTFCIKSRHVKKLQLEGRQHRSTWPTWVLFVFWVPTDAYWQILFSLVKVTRNLK 811

DB 127 LLCEGLDLPQCHLEKIQQL-----EYCRLLTAASCEPLASVLRATRAUK 168

QY 812 ELDLSGNSLSHSAVSKLTCLTRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQ-TLTEL 870

DB 169 ELTVSNNDIGEAGRVLCQGLADSACQLETLRLNCGLTTPANCKDLC-GIVASQASIREL 227

QY 871 DLSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSAPSKELDLQ 930

DB 228 DLGSNGLGDAGIAELCPGLLSPASRLKTLWLWECDDITAGCRDLQVQAKETLKELSLA 287

QY 931 QNNLDDVGVRLICEGLRHPACK-----LIRLGLDQTT 962

DB 288 GNKLDEGARLLCESLLQPGQLESWLKWSCTFTAACCOHVSLMLTQNKHLEQLSSNK 347

QY 963 LSENRQEL-RALEQKPOLLIIFSRKPSVMTPEGLDTGEMSN-TSSLKRQRRLGSRRA 1020

DB 348 LCDSGIQELCQALSQPGTTLRVLC-----LGDCEVTNSGCSSL-----A 386

QY 1021 ASHVAQANIKLLDVS 1035

DB 387 SLLANRSURELDLS 401

RESULT 7

A48843 MHC class II transactivator - human

N;Alternate names: CIITA

C;Species: Homo sapiens (man)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999

C;Accession: A48843

R;Steinmies, V.; Otten, L.A.; Zufferey, M.; Mach, B.

Cell 75, 133-146, 1993

A;Title: Complementatation cloning of an MHC class II transactivator mutated in hereditary

A;Reference number: A48843; MUID:94006536; PMID:8402893

A;Accession: A48843

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1130 <STE>

A;Cross-references: GB:X74301; NID:G414112; PIDN:CAAS2354.1; PID:G414113

Query Match 4.6%; Score 345; DB 2; Length 1130;

Best Local Similarity 23.2%; Pred. No. 4e-14;

Matches 256; Conservative 134; Mismatches 430; Indels 282; Gaps 49;

QY 13 LEFLKE-----ELKEFOLLANKAHSSSGETPA-----OPKSTSMSEVAS 55

DB 111 LEGLSKDIFKHIGPDEVIGESMEMPAAEVGQSKRPPPEELPADLKHKWPAAEPPTVVTGS 170

QY 56 YLVAQYQORANDLALHTWEQGLRSLC---AQAEAGHSPSPSPSPSEPHLGSPSP 111

DB 171 LLVGVPD-----CSTLPCLPLPALFNPQBSPAGQMRLEKTDQIPMPFSSSSUSCLNLP 223

QY 112 TSTAVLMPWIHELPACTCGSERRV-----LRQLPDTSGRRWRISASLLY 157

DB 224 EGPIQFVTITSLPHGLWQISEAGTGVSSIFIVHGEVPAQSQVPPPSG-----FTV 274

QY 158 QALPSPDPHSPQSPQSNAPTSTAVLGSWGSPQSPQSLAPR---EQEAPGTQMPDLETSG 213

DB 275 HGLTPSPDR---PGSTSPAPSATDL---PSMPEPALTSRANMTEHKTSPTQCP---AAG 325

QY 214 IYITEIREREREKSEKGRPPHAAVVGTPPQATSLQPHHHHWPSPVRSLSCTWPWKNE 273

DB 326 -----EVSNN-LPKWPEPV---EQFYRSLQDTYGA-EPAGPDGI----- 359

QY 274 FNQKFTQLLLQRPSPRSDPLVKR--SWPDVVEN--RGHLIEIRDLFGCLDQEPRI 329

DB 360 ----LVEVDLQVQARLERSSSKLERELATPDWAERQLAQGGLAELVLLAAKERRRRTRV 415

QY 330 VILQGAAGIGKSTLARQVKEAWGRGQLYGRFQHVYFSCRELAQ-SKVSLAEILGKDG 388

DB 416 IAVLKGAGQSKSYWAGAVSRWACGRL--POYDFVFSVPCHLNRPDGAAYGLQDLFLSLG 473

QY 389 TATPAPI-----RQILSRPERLLFILDGVDEFGWVLOEFPSSSELCLHWSQPQAD---- 437

DB 474 ---POPLVAADDEVFSHILKRPDRVLLILDAAFE-----LEAQDGLHSTCTGPAPAEPSL 525

```
QY 438 -ALLGSLGKTLPEASFIITARTTALONLPSLEQA-RWVEVLGSPSSSRKEYFYRYPT 495
D 526 RGLLAGLFOKKLIRGCTLLITARPG--RLVQSLSKADALFELSGSPMEQAQAYVMRYFE 583
QY 496 D-----BROAIRAPRLVYSKNEKALCLVPMVSWLACTCLMOQMKREKLTLSKTTTLC 551
D 584 SSGMTEHQD-RALTURDRPRLLSHSHSPTLCRAVCQLSEALLELGEDAKLPS-TUTGLY 641
QY 552 LHYLAQALQAPLGPQLRDLCSLAAB-GIWQKTL----FSPDDLKRGHGLGALISTFLK 606
D 642 VGLLGRAALDSPG-LAELAKLAWELGRRHQSTLQEDQFSPADVR-----TWANA 691
QY 607 MGILQHP-----IPLSY-SFIHLCF-OEFFAAMSYYLEDE-----KGRGHSNCI 650
D 692 KGLVQHPRAAESLAPFPFLQCFLCALWALSGEIKDKELPQYLALTPRKRPYDNWL 751
QY 651 IDEKTEAVGIHGLFGASTTRFLLGLSDEG-----EREM 686
D 752 EGVPRFLA-----GLIFQPPARCGLGALLGFSAAASVDRKQVLARVYLKRLQPGTTLRQ 806
QY 687 ENIFHC-----RLSQGRNLMQWVP-----SIQLLLQPHSLB----- 717
D 807 LELHCAHEABEAGIWHQVQELPGRSLFLGTTLTPDAHVGLKALEAAGQDFSLDRST 866
QY 718 -----SLHLYETRNKTLFQVMAHFEEMGCVETDM-----ELLVCTFCIKP 760
D 867 GICPSGLSLVGLSCV--TRFAALSDTVALWBSLRHQGETKLLQAAEBKFTIEPPKAKS 924
QY 761 SRHVKL-QLIEGRQHRSTWSPMVL-----FRWVPVTDAY-WQILFSLVKVTR 808
D 925 LKQVEDGLKLVQTRTRESSDAGELPAVRDLKLEFALGPVSPQAPFPKLVRILTAFS 984
QY 809 NLKELD--LSGNSLSHSAVSKLCTLRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQ 865
D 985 SIQHLDDALSENKIGDEGVQSLSATF--POL-----K 1015
QY 866 TLTELDLSFNVLTDAGAKHLQORLPQSCKLQRLQVSCGLTSDCCQDLASVLSASPSLK 925
D 1016 SLETLNLSQNNITDLGAYKLAELPSLAASLLRLSLVNNICIDVGAESLARVLDPMVSLR 1075
QY 926 ELDLQONLDDVGVRLICEGLR 947
D 1076 VMDVQYNKFTAGAQAQLAASLR 1097

RESULT 8
T31668
hypoetical protein COSI.5 - sea squirt (Ciona intestinalis)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T31668
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: EMBL:Z81050
A:Accession: T31668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1075 <IR>
A:Cross-references: EMBL:Z80904; PIDN: CAB02589.1
C:Genetics:
A:introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1

Query Match 2.7%; Score 206; DB 2; Length 1075;
Best Local Similarity 18.5%; Pred. No. 4.8e-05;
Matches 209; Conservative 152; Mismatches 347; Indels 422; Gaps 49;

QY 301 PDYVEENRGLHBIRODLFGGLDT-----OEPRIVLQGAAGICKST 342
D 37 PEPAPVNVQLVTDQLFDKALENAQRYTEDESAEYAKVIERHANTVWVGPPGVGKTT 96
QY 343 LARQVKEMARGQLYGRFOHFYFSCRELAQSKVSLAELIGKDGTA-----TPAP 394
D 97 LKQWVKQILKHEFLPDT-EYIFFINVKOIDFNKMTLLEFLTNSGRVKNVYTERESKAL 155
```

RESULT 9

T31665

hypoetical protein COSI.4 - sea squirt (Ciona intestinalis)

```
QY 395 IRQILSRPERLLFIIDGVDEPGWVLQEPSSSELCLHWSQPQPADALLGSLGKTLPEASF 454
D 156 ITFLEHNNPVAIF-FDGLDEASTNEFARIPHIKCDGSKPVD-IMKNLNLTLPLPKAKI 213
QY 455 LITARTTALONLPSLEQARWVEVLGSPSSSRKEYFYRYFTDERQAIRAFRLVKSNEKLW 514
D 214 VVTSLHQMYKLHPYRPTSFIVELGLLEAAKNGNLTQLCGEKYPAIK--KILDQOPNLA 271
QY 515 ALCLVPWVSWLACTCLMOQMKREKLTLS-----KTTTTLCLHYLAQALQAPLGPQLRD 570
D 272 HLCYLPINFIIVFCLLN-EGSDIKTMTQVLI-FMTRFVELSHLKGVPLOKVGAEWK 330
QY 571 LCSLAAGIWOQKTLFSPDDLKRGHGLDGAISTF-----LKMGILOBHPILPSYS 620
D 331 LARLAYKGLQORLKFVEKTFDDVDKLABEMVTNFFHTYVDKSGIRMKILEGN--KRSY- 387
QY 621 FTHLCFOBPFAAMSIVL----- 637
D 388 FTHLIWQEFYAAVYLMFLFVSYPREFQLKPIFKDAQWKRVVGMFGICNPPAYKQLKLP 447
QY 638 -----EDEK-----GRG----- 644
D 448 ATMWDYBEKKELMVMPMMSLSARGEDLIRRFGLWHEYNDDSSKXEPEDYLPVGLKMD 507
QY 645 --KHSNCIIDLKTELEY-----GIHGLFGASTTRFLLG-- 676
D 508 APKHLSEVKDLVYALKSFTPKHLRLDSYETTTTTEVLETLRLRGVHGT-TTITRFRVINNI 566
QY 677 -----LLSDGEREMENI-----FHCRLSQGRNLMQWVPSLQLLOPHSLES 719
D 567 EMKDSIMELLHLLHLDAMBELRPIRVNTNLPYMERLSNAIN--QRSNKIQLVLIKH--KL 621
QY 720 HCLYETRNKTLFQVMAHFEEM--GMCVETDMEILLVCTFCIKFSHVVKLQLIEGRQHR 776
D 622 HDYDVK--YLAGCLGNISLLYMMGTDISSQ-----CSVLKQAIQOLPSIQ-- 666
QY 777 STWSPMTVVLPFRVPVTDAYWOILFSLVKVTRN-----LKELDLSGNSLSHSAVSLCK 830
D 667 -----VHQLYPDILSTYLVNARNWISNMFVSEVDLVPEALN----- 702
QY 831 TLRRPCLLETURLAGCGLTAEDCKDLAPGLRANQTLTDLDSFNVLTDAGAKHLQORLR 890
D 703 -----LGNGL-----KDSKQOPRRNEELCSQDVSVPAPTQVNNHYC--N 740
QY 891 QPSCKLQRLQLVSCGLTSDCCDLASVLSASPSLKELDLQONLDDV----- 937
D 741 APYHQVRFL-----TQAFILF--MINSNLHPNGLDQVEKIRWLEDESEYLE 784
QY 938 GVRLLCEGLRHPA-----CKLI-----RLGLD 959
D 785 TPKLQCE-----PASLTFKEQVTKLQTHVVDKETIRCKLLHSKDGITWEETKTKLEFS 840
QY 960 QTTLS-----DEMQRERLALRQEKPKQLLIFSRRKPSVMTPTGLD--TGEMS 1004
D 841 DNFISFQTFNYSWVKVIFEFIEKTLAFYKVLKSLLYQAGDVWVKWCWNLDDVATGE-- 898
QY 1005 NSTSSILKQRLGSEARAASHVAQANLKLVDVKIFPIAEIABESSPEVVPVLCVPS-- 1061
D 899 ---NELKENHFTQETWT--IPLSNDLILCLEK-----HENEAEVINI----IPSGKI 941
QY 1062 -PASQ-----GDLHTKPLGTDDDQWGPVTAEDVVDKEKLYRVHFFPVAGSY 1108
D 942 IPANQLNNSYCCNKKFKVDKHS---TNE-----VHLIAKAK-----CGSF 978
QY 1109 RWPNTGLCFWMBREAVTVEIEFCVWDQ---FLGEINPQHSMMWVAGPLLDIK 1155
D 979 RW-DDDFCF-----PLPITHAVSDQQOPTQPSINP--TVIVTGPPTTIR 1019
```


Db 257 EAAKVAELIPSTEKLVLFHNNMTDEGAFAIAEVRKSRPALEDFRC---SSTRVSGE 313
 QY 733 QVMAHFEEMCMCVETDMELLVCTFCIKFSRHKVKKLQIEGRQHRSTWSTMVVLFRWVPV 792
 Db 314 GGVALAEALGACT-----HUKKIDL----- 333
 QY 793 TDAYMOI-----LFSVLKVTNLIKELDLSGNSLSHSAVSKLCTLRPRCLLETIRLAGC 847
 Db 334 RDNMFGEAGVALSKVIPFADLTETIYLSVNLNEDDGAALANALAKESAPSETLDMAGN 393
 QY 848 GITAEDCKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLQRPCKLQRLQVSCGLT 907
 Db 394 DITAKATVSVAECISSKQPLTKNLSENELKDBGA-----GL- 430
 QY 908 SDCCQDLASVLSAPSLSKELDLQNNLDDVGRLLCEG-LRHPACKLIRLGLDQTLSD 966
 Db 431 -----ISKALREGGOLSEVDLSTNLITWGAKLAEAVVQKPGFKL--LNINANFISDE 482
 QY 967 MRQELRALQEKPOLLIFFSRKPSVMTPTTEGLDTGEMSNSTS 1008
 Db 483 GIDELKXIFKSPDML-----GELDENDPEGEDVDDEAEDDS 519

RESULT 13
 JC7765
 Mitotic spindle associated protein, MAP126 - human
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 R:Chang, M.S.; Huang, C.J.; Chen, M.L.; Chen, S.T.; Fan, C.C.; Chu, J.M.; Lin, W.C.; Yan
 Biochem. Biophys. Res. Commun. 287, 116-121, 2001
 A:Title: Cloning and characterization of hMAP126, a new member of mitotic spindle-associated
 A:Reference number: JC7765; PMID:11549262
 A:Contents: Testis
 A:Accession: JC7765
 A:Molecule type: mRNA
 A:Residues: 1-1120 <CHA>
 A:Cross-references: GB:AF345347
 C:Comment: This protein, as a nuclear protein and a new member of kinesin family, which
 and dynamic regulation of mitotic spindles and may provide an outward force to push the
 C:Genetics:
 A:Gene: hmap126
 A:Map position: 17q11-12

Query Match 2.2%; Score 166.5; DB 2; Length 1120;
 Best Local Similarity 18.8%; Pred. No. 0.02;
 Matches 217; Conservative 138; Mismatches 391; Indels 411; Gaps 50;

QY 148 WR--EISALLYQALPSPDHSPQSPESNPAPTSTAVLGSWSPPOPSLAPR-----EQE 200
 Db 2 WRVKLSLSLSPSPQTKPSMRTPLELTLPQALTSK-RSPACSSLTPLSKLGLQE 60
 QY 201 APQTOWPLD-----ETSGIYYTE-----IREREREKSEKGRPPAAVVGTPPQAHTSLQ 249
 Db 61 GSNSSPVDVFNKRDTLSEHFSHSSKWLCTQCHESDEQPLDPIQISSTRTKTSSEAVD 120
 QY 250 PHHPHPEPSVRESLCSTWPKNEDFNQFTQLLQRPSPQDPLVKSFWPDYVENRG 309
 Db 121 PLG-----NYMVKTVLVSPLGQQQDM----- 143
 QY 310 HLIEIRDLFGPGLDQ-EPRIVILQAAGTIGKSTLARQVKEAWGRQLYGDGRFQHFVYFS 368
 Db 144 -----IFEARLDTAETNSISLNGP--LRDTDVREEV-----APCMGDRF----- 182
 QY 369 CRELAQKVSVAELIKGDKGTATPARIQILSRPERLLFDGVDPGFWLQEPSELSC- 427
 Db 183 -----SEVAAVSE-----KEIFQ-----ESPSHLLSESPNPSC 211
 QY 428 --LHWSQOPADALLOSILGKLTILPEASFLITARTTAL--ONLIPSLQA-----RWREV 478
 Db 212 EQLHCKES-----LSRTEAVREDLVPSSENAFLPSSVLW----- 247
 QY 479 LGFSESSRKEYFYRYFTDRQAI-----RAPR-----LVKSKNELWALC 517

Db 248 LSPSTALAADFRVNHVDPEEIVEHGAMEERMPPTHPKESETEDQALVSSVEDILSTC 307
 QY 518 LVP-----WVSWLACTCLMQQMKRKEKLTLSKTTTT 549
 Db 308 LTPNLVEMESQAPGPAVEDVGRILGSDTESMWSPLA-----WLEKGVNTS 353
 QY 550 LCLHLAQAQAQPLGPQRLDLC-----SLAAEGIMQKTLFSPD-----DLRKHGL 596
 Db 354 VMLNLRSLGSL--LPSMLRDAAGTTTFTSCSVGTW-----FTPSAQEKSTNTSOTGL 405
 QY 597 DGAIISTFKMGILQEHPIPLSYSEFIH-----LCFQEFFAAMSIVLEDEKGR----- 643
 Db 406 VTKHSTSETQLLCGRPPDLTALSRHDLNLLSSLVIVEF---LSRQLRDWKSQALVP 462
 QY 644 -----GKHSNCIIDLKLEAYGHLGFGASTTRFLGLLSDGEGERENEFH 691
 Db 463 HPETQDSSTQTDTSHTSGITNKLOHLKESH-----EMGO----- 495
 QY 692 CRLSQGRNLMO-WVPSLOLLQPHSLHSLHCLYETRNKTLFTQ----- 733
 Db 496 -ALQOARNVMSWVLSKELI---SLHLSLHLLEEDKTTVNOESRRRAETLVCCCFDLK 551
 QY 734 -----VMAHFEEMCMCVETDMELLVCTFCIKFSRHKVKKLO--LIEGRQHR 777
 Db 552 KLRAKLQSLKAEREAREHREEMALRGDAAEIVLEAFCAHASQRIQSLEQDLASMRFRG 611
 QY 778 TWSPTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCTLRPRC 837
 Db 612 LKDAQTQL-----VGLHAKQEBELVQQTSLTSLTQ--DWRSMLQDYTTWTAL---LSRSRQ 664
 QY 838 LLETURLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAK--AKHLQRLQRPCK 895
 Db 665 LTKLT-----VKSQALQERDVAIEKEQSVSVLSQVSALECKGQ 707
 QY 896 LQRLQVSCGLTSDCCQDLASVLSAPSLSKELDLQNN-LDDVGRV--LLCE----- 944
 Db 708 TEQLELENIRLATDLRAQQLANMDSQKELQSHQTHCAQDLAMKDLCLQTSNEEQ 767
 QY 945 -----GLRHPACKLIRLGLDQTLSDMRQELRALE-----OE 977
 Db 768 AAQCVKEEMALKHMQAELOQ---QOAVLAKEVRDLKETLEFADQENQVAHLELQGVQCOL 824
 QY 978 KPQLLIFFSR-----KPSVMTPTTEGLDTGEMSNSTSLSKQRLGSRASASHVAQANLK 1030
 Db 825 KTLLEVLERSLOENLKDVTENLTAKLASTIADNOEQDLKTRQYSOKLGLLTQ----- 880
 QY 1031 LLDVSKIFPIAIEAESPEVVPVELLCVPSPASQGLDHTKPLGTDGDFWGTGVPATEV 1090
 Db 881 -LQSLTLFLQTKLKEKTEQETLLILLSTACPT-----QEHPLPNDRTF---LGSILTAV 929
 QY 1091 VDKEKVLVRVHPVNGS 1107
 Db 930 ADEEPESTPV--PLIGS 944

RESULT 14

A45841
 T-complex-associated-testes-expressed-1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
 R:Accession: A45841
 R:Sarvetnick, N.; Tsai, J.Y.; Fox, H.; Pilder, S.H.; Silver, L.
 Immunogenetics 31, 283-284, 1990
 A:Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
 A:Reference number: A45841
 A:Accession: A45841
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-506 <SAR>
 A:Cross-references: GB:M28821; NID:g201909; PIDN:AAA40406.1; PID:g201910

Query Match 2.1%; Score 159.5; DB 2; Length 506;

Db 918 SP2EI 922

Search completed: January 29, 2004, 13:49:45
Job time : 40.197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 21.9982 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWGRACYLEFLKKEE.....HLIMELWEKSGKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7502	99.6	1473	1 NALI_HUMAN	Q9C000 homo sapien
2	1382.5	18.4	1062	1 PYAT_HUMAN	P59046 homo sapien
3	1279	17.0	1034	1 CISI_HUMAN	Q9P200 homo sapien
4	1254.5	16.7	1033	1 CISI_MOUSE	Q8R4B8 mus musculus
5	1037	13.8	994	1 NALA_HUMAN	Q96nm2 homo sapien
6	983	13.0	1200	1 MATE_HUMAN	P59047 homo sapien
7	948	12.6	1062	1 NALA_HUMAN	Q9nx02 homo sapien
8	908.5	12.1	980	1 PYA3_HUMAN	Q8wx94 homo sapien
9	901.5	12.0	892	1 PYA5_HUMAN	P59044 homo sapien
10	895	11.9	854	1 PYA5_RAT	Q63035 rattus norv
11	882	11.7	1111	1 MATE_MOUSE	Q9rlm5 mus musculus
12	868	11.5	843	1 PYA5_MOUSE	Q9lwe2 mus musculus
13	711	9.4	1033	1 PYA6_HUMAN	P59045 homo sapien
14	668	8.9	431	1 CAR8_HUMAN	Q9y2g2 homo sapien
15	393.5	5.2	1040	1 CARF_HUMAN	Q9hc29 homo sapien
16	392	5.2	1020	1 CARF_MOUSE	Q8k3z0 mus musculus
17	388.5	5.2	953	1 CAR4_HUMAN	Q9y239 homo sapien
18	361.5	4.8	456	1 RINI_RAT	P29315 rattus norv
19	358	4.8	460	1 RINI_HUMAN	P13489 homo sapien
20	354.5	4.7	456	1 RINI_PIG	P10775 sus scrofa
21	351	4.7	953	1 CAR4_MOUSE	Q8bhb0 mus musculus
22	345	4.6	1130	1 C2TA_HUMAN	P33076 homo sapien
23	318	4.2	1155	1 C2TA_MOUSE	P79621 mus musculus
24	253.5	3.4	193	1 ASC_MOUSE	Q9epb4 mus musculus
25	250.5	3.3	195	1 ASC_HUMAN	Q9ulz3 homo sapien
26	215	2.9	1024	1 CARC_HUMAN	Q9npb4 homo sapien
27	212.5	2.8	1403	1 BIRE_MOUSE	Q9r016 mus musculus
28	200	2.7	1403	1 BIRE_MOUSE	Q9jib6 mus musculus
29	199	2.6	1402	1 BIRG_MOUSE	Q9jib3 mus musculus
30	198	2.6	1403	1 BIRI_HUMAN	Q13075 homo sapien
31	196.5	2.6	1447	1 BIRI_MOUSE	Q9quk4 mus musculus
32	196	2.6	1403	1 BIRA_MOUSE	Q9gwks mus musculus
33	161	2.1	483	1 YAX2_HUMAN	Q95522 homo sapien

RESULT 1									
NALI_HUMAN									
ID	NALI_HUMAN	STANDARD;	PRT;	1473	AA.				
AC	Q9C000; Q9BZ28; Q9BZ29; Q9HAV8; Q9UFT4; Q9Y2E0;								
DT	16-OCT-2001 (Rel. 40, Created)								
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DT	15-SEP-2003 (Rel. 42, Last annotation update)								
DE	NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-forming ced-4-like apoptosis protein) (Nucleotide-binding domain and caspase recruitment domain) (Caspase recruitment domain protein 7).								
DE	GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A. (ISOFORM 2).								
RX	MEDLINE=21169419; PubMed=11270363;								
RA	Bertin J., Distefano P.S.;								
RT	"The PYRIN domain: a novel motif found in apoptosis and inflammation proteins.";								
RT	Cell Death Differ. 7:1273-1274(2000).								
RN	[2]								
RP	SEQUENCE FROM N.A. (ISOFORM 2).								
RX	MEDLINE=21148093; PubMed=11250163;								
RA	Martinon F., Hofmann K., Tschopp J.;								
RT	"The pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation.";								
RT	Curr. Biol. 11:R118-R120(2001).								
RN	[3]								
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).								
RC	TISSUE=Erythrocyte;								
RX	MEDLINE=21153743; PubMed=11076957;								
RA	Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A., Shi M.M., Vincenz C., Ward P.A.;								
RT	"Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis proteins.";								
RL	J. Biol. Chem. 276:9230-9238(2001).								
RN	[4]								
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION								
RC	TISSUE=T-cell;								
RX	MEDLINE=21153744; PubMed=1113115;								
RA	Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S., Godzik A., Reed J.C.;								
RT	"A novel enhancer of the Apaf1 apoptosome involved in cytochrome c-dependent caspase activation and apoptosis.";								
RL	J. Biol. Chem. 276:9239-9245(2001).								
RN	[5]								
RP	SEQUENCE FROM N.A. (ISOFORM 2).								
RC	TISSUE=Brain;								
RX	MEDLINE=99246063; PubMed=10231032;								
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;								
RT	"Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code								

34	158	2.1	5065	1	EPPL_HUMAN	P58107 homo sapien
35	152	2.0	3511	1	MY15_MOUSE	Q9qz24 mus musculus
36	147.5	2.0	793	1	ST5A_MOUSE	P42230 mus musculus
37	146.5	1.9	1524	1	Y133_HUMAN	Q14146 homo sapien
38	145.5	1.9	793	1	ST5A_RAT	Q62771 rattus norv
39	144	1.9	2300	1	CYAA_NEUCR	Q01631 neurospora
40	142.5	1.9	4303	1	PKD1_HUMAN	P98161 homo sapien
41	142	1.9	1426	1	NPH4_HUMAN	O75161 homo sapien
42	142	1.9	1839	1	CYAA_SACKL	P23466 saccharomyc
43	141.5	1.9	1021	1	PSKR_DAUCA	Q81pb4 daucus caro
44	140	1.9	794	1	ST5A_BOVIN	Q95115 bos taurus
45	138.5	1.8	794	1	ST5A_HUMAN	P42229 homo sapien

ALIGNMENTS

RT for large proteins, in vitro.";

RL DNA Res. 6:63-70(1999).

RN [6]

RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).

RA TISSUE=Uterus;

RC Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wienann S.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Able to form cytoplasmic structures termed death

CC effector filaments. Enhances APAF1 and cytochrome c-dependent

CC activation of pro-caspase-9 and consecutive apoptosis. Seems to

CC bind ATP.

CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9

CC and with APAF1 in a cytochrome c-inducible way leading to the

CC formation of an apoptosome. This interaction may be ATP-dependent.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Name=1; Synonyms=NAC beta, DEFCAP-L;

CC IsoId=Q9C000-1; Sequence=Displayed;

CC Name=2; Synonyms=NAC alpha, DEFCAP-S;

CC IsoId=Q9C000-2; Sequence=VSP_004327;

CC Name=3; Synonyms=NAC gamma;

CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;

CC Name=4; Synonyms=NAC delta;

CC IsoId=Q9C000-4; Sequence=VSP_004326;

CC -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are

CC expressed in peripheral blood leukocytes, chronic myelogenous

CC leukemia cell line K-562, followed by thymus, spleen and heart.

CC Also detected in lung, placenta, small intestine, colon, kidney,

CC liver and muscle.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 1 CARD domain.

CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF298548; AAG15254.1; -

CC EMBL; AF310105; AAG30288.1; -

CC EMBL; AF229059; AAK00748.1; -

CC EMBL; AF229060; AAK00749.1; -

CC EMBL; AF229061; AAK00750.1; -

CC EMBL; AF229062; AAK00751.1; -

CC EMBL; AB023143; BAA76770.1; -

CC EMBL; AL117470; CAB55945.1; -

CC PIR; T17255; T17255.

CC HSP; P13489; 1A4Y.

CC MIM; 606636; -

CC GO; GO:0005622; C:intracellular; IC.

CC GO; GO:0016506; F:apoptosis activator activity; NAS.

CC GO; GO:0008656; F:caspase activator activity; NAS.

CC GO; GO:0019899; F:enzyme binding activity; IPI.

CC GO; GO:0006919; P:caspase activation; NAS.

CC GO; GO:0006917; P:induction of apoptosis; NAS.

CC InterPro; IPR001315; CARD.

CC InterPro; IPR000767; Disease_resist.

CC InterPro; IPR001611; LRR.

CC InterPro; IPR007091; LRR_RNinh.

CC InterPro; IPR007111; NACHT_NPase.

CC InterPro; IPR004020; PAAD_DAPIN_dom.

CC Pfam; PF00560; LRR; 2.

CC Pfam; PF02758; PAAD_DAPIN; 1.

CC PRINTS; PR00364; DISEASERSIST.

CC PROSITE; PS02009; CARD; 1.

CC PROSITE; PS00824; DAPIN; 1.

CC PROSITE; PS00837; NACHT; 1.

CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;

Query Match	Score	7502;	DB 1;	Length	1473;
Best Local Similarity	97.0%;	Pred. No. 0;			
Matches 1429;	Conservative	0;	Mismatches	0;	Indels 44;
					Gaps 1;
QY 1	MAGGAWGLACVLRFLKKEELKEFQLLLANKAHSSSSGETPAQPEKTSGMVEASVLAQ	60			
Db 1	MAGGAWGLACVLRFLKKEELKEFQLLLANKAHSSSSGETPAQPEKTSGMVEASVLAQ	60			
QY 61	YGEQRAWDLALHTWEQMLRSLCAQAQAGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW	120			
Db 61	YGEQRAWDLALHTWEQMLRSLCAQAQAGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW	120			
QY 121	IHELPAQTQGSRRVLRQLPDTSGRRWRREISASLLYQALFSSPDHSPSESNAPTST	180			
Db 121	IHELPAQTQGSRRVLRQLPDTSGRRWRREISASLLYQALFSSPDHSPSESNAPTST	180			
QY 181	AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYTYEIREREREKSEKGRPPAAVVG	240			
Db 181	AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYTYEIREREREKSEKGRPPAAVVG	240			
QY 241	PPQATSLQPHHPWEPSPVRESLCSCTWPWKNEDEFNQKFTQLLLQRPSPQDPLVKRSW	300			
Db 241	PPQATSLQPHHPWEPSPVRESLCSCTWPWKNEDEFNQKFTQLLLQRPSPQDPLVKRSW	300			
QY 301	PDYVEENRGLHIEIRDLFGPGLDTQEPRIVLQGAAGIKGKSTLARQVKEANGRGOLYGD	360			
Db 301	PDYVEENRGLHIEIRDLFGPGLDTQEPRIVLQGAAGIKGKSTLARQVKEANGRGOLYGD	360			
QY 361	FQHVYFSCRELAQSKVVSLELIGKDGATAPATPIQLSRPERLLFLDGDVDEPGWLQ	420			
Db 361	FQHVYFSCRELAQSKVVSLELIGKDGATAPATPIQLSRPERLLFLDGDVDEPGWLQ	420			
QY 421	EPSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSLQEARVVEVLG	480			
Db 421	EPSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSLQEARVVEVLG	480			
QY 481	FSESRKEYFYRYFTDERQAIARFRLVKSNKELWALCLVPVWSMLACTCLQOQMKRKEKL	540			
Db 481	FSESRKEYFYRYFTDERQAIARFRLVKSNKELWALCLVPVWSMLACTCLQOQMKRKEKL	540			
QY 541	TLTSKTTTTTCLHVLQAQALQAPLGPQRDICSAAEGIWOKKTLFSPDDLRKHGLDGA	600			
Db 541	TLTSKTTTTTCLHVLQAQALQAPLGPQRDICSAAEGIWOKKTLFSPDDLRKHGLDGA	600			

```
QY 601 ISTFLKMGILQHEPIPLSYFHLFCQEPFAAMSVYVLEDEKGRGKHSNCIIDLKXTLEAY 660
Db 601 ISTFLKMGILQHEPIPLSYFHLFCQEPFAAMSVYVLEDEKGRGKHSNCIIDLKXTLEAY 660
QY 661 GHGLFGASTTFLGLLSDEGEREMENIFHCRLSQGRNLMQVPSLQLLLOPHSLESILH 720
Db 661 GHGLFGASTTFLGLLSDEGEREMENIFHCRLSQGRNLMQVPSLQLLLOPHSLESILH 720
QY 721 CLYETRNTKFTLVQMAHFEMGCVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Db 721 CLYETRNTKFTLVQMAHFEMGCVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
QY 781 PTMVVLFWRVPVTDYQWQLFSLVKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
Db 781 PTMVVLFWRVPVTDYQWQLFSLVKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
QY 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLQPSCKLQRLQ 900
Db 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLQPSCKLQRLQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRLGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRLGLDQ 960
QY 961 TTLSDEMQRLEALQEKPOLLIIFSRKPSVMTPTGLDTGEMSNSTSLKRLQSGSERA 1020
Db 961 TTLSDEMQRLEALQEKPOLLIIFSRKPSVMTPTGLDTGEMSNSTSLKRLQSGSERA 1020
QY 1021 ASHVAQANIKLIDVSKIFPIAETAESSPEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1080
Db 1021 ASHVAQANIKLIDVSKIFPIAETAESSPEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1080
QY 1081 GPTGPVATEVDKKNLYRHPVAGSYRWPNTGLCFVNRVAVTVIEFCVMDQFLGEIN 1140
Db 1081 GPTGPVATEVDKKNLYRHPVAGSYRWPNTGLCFVNRVAVTVIEFCVMDQFLGEIN 1140
QY 1141 PQHSMVAGPLDIDKAEPAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Db 1141 PQHSMVAGPLDIDKAEPAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
QY 1201 ELHHIVLENPSPGLGVLLKMHNLRFPTVTSVLLYRHPVEEVTFLYIPSDCSIR 1260
Db 1201 ELHHIVLENPSPGLGVLLKMHNLRFPTVTSVLLYRHPVEEVTFLYIPSDCSIR 1260
QY 1261 -----XELCYRSPGSDQLF 1276
Db 1261 KAIDDEMKQFQVRIHKPPPLTYMGCRYTSGSGSGMLELPLKLELCYRSPGSDQLF 1320
QY 1277 SEFYVGHLSGIRLQVKKDETLVWEALVKGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
Db 1321 SEFYVGHLSGIRLQVKKDETLVWEALVKGDLMPATTLIPPARIAVPSPLDAPQLLH 1380
QY 1337 FVDQYREQLIARTVSEVVLKHLGOVLQSOEQLVLAENTRPSQMRKLFSLQSQWDRKC 1396
Db 1381 FVDQYREQLIARTVSEVVLKHLGOVLQSOEQLVLAENTRPSQMRKLFSLQSQWDRKC 1440
QY 1397 KDGLYQALKETHPHLMELWEKSGKGLPLSS 1429
Db 1441 KDGLYQALKETHPHLMELWEKSGKGLPLSS 1473

RESULT 2
PYA7_HUMAN
ID PYA7_HUMAN STANDARD; PRT; 1062 AA.
AC P59046;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE PYRIN-containing APAFI-like protein 7 (Monarch-1).
GN PYPAF7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
```

```
NCBI_TaxID=9606;
(1)
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.,
RT "PYPAF7, a novel PYRIN-containing APAFI-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing."
RL J. Biol. Chem. 277:29874-29880(2002).
(2)
SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Lymphoma;
RA Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
(3)
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Leukocyte;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NP-kappa-B via IKK.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=1;
CC IsoId=P59046-1; Sequences=Displayed;
CC Name=2; Synonyms=11;
CC IsoId=P59046-2; Sequences=VSP_005524;
CC Name=3; Synonyms=111;
CC IsoId=P59046-3; Sequences=VSP_005523;
CC -!- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
CC predominantly in eosinophils and granulocytes, and at lower levels
CC in monocytes.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```


phenotype including recurrent fever, cold sensitivity, sensorineural deafness, and AA amyloidosis.";

[4] Arthritis Rheum. 46:2445-2452(2002).

SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).

TISSUE=Blood;

RX MEDLINE=20499367; PubMed=11042152;

RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,

RA Zhang Q.-H., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,

RA Zhang Y., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.,

RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.,

RT "Cloning and functional analysis of cDNAs with open reading frames for

RT 300 previously undefined genes expressed in CD34+ hematopoietic

RT stem/progenitor cells.";

RL Genome Res. 10:1546-1560(2000).

[5]

RN VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND

RP ARG-569, AND VARIANT FCAS/MWS TRP-260.

RX MEDLINE=21987640; PubMed=11992256;

RA Dode C., Le Du N., Cuisset L., Letourneur P., Berthelot J.-M.,

RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,

RA Granel B., Frances C., Garcier F., Edery P., Boulinguez S.,

RA Domergues J.-P., Delpech M., Grateau G.;

RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome

RT and familial cold urticaria: a novel mutation underlies both

RT syndromes.";

RL Am. J. Hum. Genet. 70:1498-1506(2002).

[6]

RN VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND

RP THR-662, AND TISSUE SPECIFICITY.

RX MEDLINE=22062556; PubMed=12032915;

RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,

RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;

RT "Chronic infantile neurological cutaneous and articular syndrome is

RT caused by mutations in CIAS1, a gene highly expressed in

RT polymorphonuclear cells and chondrocytes.";

RL Am. J. Hum. Genet. 71:198-203(2002).

CC -!- FUNCTION: May function as a potential inducer of apoptosis.

CC Interacts selectively with apoptosis-associated specklike protein

CC containing a CARD domain (ASC). This complex may function as an

CC upstream activator of NF-kappaB signaling.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Name=2;

CC IsoId=Q96P20-1; Sequence=Displayed;

CC Name=1;

CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;

CC Name=3;

CC IsoId=Q96P20-3; Sequence=VSP_005519;

CC -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly

CC expressed in polymorphonuclear cells, undetectable or expressed

CC at a lower magnitude in B and T lymphoblasts, respectively. High

CC level of expression detected in chondrocytes. Low or no expression

CC in the other tissues tested.

CC -!- DISEASE: Defects in CIAS1 are a cause of familial cold

CC articularia. FCAS is rare autosomal dominant systemic inflammatory

CC disease characterized by episodes of rash, arthralgia, fever and

CC conjunctivitis after generalized exposure to cold.

CC -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome

CC (MWS), a rare autosomal dominant fever syndrome with episodic

CC urticaria, arthralgia, amyloidosis and progressive sensorineural

CC deafness.

CC -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile

CC neurologic cutaneous and articular syndrome (CINCA), also known as

CC 'neonatal onset multisystem inflammatory disease,' or NOMID, a

CC rare congenital inflammatory disorder characterized by a triad of

CC neonatal onset of cutaneous symptoms, chronic meningitis, and

CC joint manifestations with recurrent fever and inflammation.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts

CC in positions 893, 918 and 926.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF410477; AAL33908.1; -

DR EMBL; AF427617; AAL33911.1; -

DR EMBL; AY051117; AAL12497.1; JOINED.

DR EMBL; AY051112; AAL12497.1; JOINED.

DR EMBL; AY051113; AAL12497.1; JOINED.

DR EMBL; AY051114; AAL12497.1; JOINED.

DR EMBL; AY051115; AAL12497.1; JOINED.

DR EMBL; AY051116; AAL12497.1; JOINED.

DR EMBL; AY056059; AAL12497.1; JOINED.

DR EMBL; AY056060; AAL12497.1; JOINED.

DR EMBL; AY051117; AAL12498.1; -

DR EMBL; AY051112; AAL12498.1; JOINED.

DR EMBL; AY051113; AAL12498.1; JOINED.

DR EMBL; AY051114; AAL12498.1; JOINED.

DR EMBL; AY051115; AAL12498.1; JOINED.

DR EMBL; AY051116; AAL12498.1; JOINED.

DR EMBL; AF420469; AAL65136.1; -

DR EMBL; AF468522; AAL78632.1; -

DR EMBL; AF492033; AAM14669.1; ALT_INIT.

DR EMBL; AF418985; AAL14640.2; ALT_INIT.

DR EMBL; AF04176; AAC39910.1; ALT_FRAME.

DR Genem; HGNC:16400; CIAS1.

DR MIM; 606416; -

DR MIM; 120100; -

DR MIM; 191900; -

DR MIM; 607115; -

DR GO; GO:0046506; F:apoptosis activator activity; NAS.

DR GO; GO:0006917; P:induction of apoptosis; NAS.

DR GO; GO:0006954; P:inflammatory response; IMP.

DR GO; GO:0007165; P:signal transduction; NAS.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007091; LRR_RNinh.

DR InterPro; IPR003590; LRR_RNinh_sub.

DR InterPro; IPR007111; NACHT_NTPase.

DR InterPro; IPR004020; PAAD_DAPIN_dom.

DR Pfam; PF00560; LRR; 2.

DR SMART; SM00368; LRR_R1; 3.

DR PROSITE; PS50824; DAPIN; 1.

DR PROSITE; PS50837; NACHT; 1.

DR Apoptosis; Repeat; Leucine-rich repeat; Alternative splicing;

DR Disease mutation; Deafness.

FT DOMAIN 1 91

FT DAPIN.

FT DOMAIN 218 534

FT NACHT.

FT REPEAT 738 761

FT LRR 1.

FT REPEAT 795 818

FT LRR 2.

FT REPEAT 852 875

FT LRR 3.

FT REPEAT 881 904

FT LRR 4.

FT REPEAT 909 937

FT LRR 5.

FT REPEAT 939 961

FT LRR 6.

FT REPEAT 966 989

FT LRR 7.

FT POLY-GLU.

FT DOMAIN 688 695

FT Missing (in isoform 3).

FT VARSPLIC 718 1034

FT /FTID=VSP_005519.

FT VARSPLIC 719 775

FT Missing (in isoform 1).

FT /FTID=VSP_005520.

FT VARSPLIC 834 890

FT Missing (in isoform 1).

FT /FTID=VSP_005521.

FT V -> M (IN FCAS AND MWS).

FT /FTID=VAR_013227.

FT R -> W (IN FCAS AND MWS).

FT /FTID=VAR_014104.

FT D -> N (IN CINCA AND MWS).

FT /FTID=VAR_014105.

FT L -> P (IN FCAS AND MWS).

FT	VARIANT	309	309	/FTid=VAR_014124. F -> S (IN CINCA). /FTid=VAR_014106. T -> M (IN MWS). /FTid=VAR_014366. A -> V (IN MWS). /FTid=VAR_013228. H -> R (IN CINCA). /FTid=VAR_014367. T -> N (IN CINCA). /FTid=VAR_014368. A -> T (IN MWS). /FTid=VAR_014369. A -> V (IN FCAS). /FTid=VAR_013229.	17.0%; Score 1279; DB 1; Length 1034; Best Local Similarity 33.0%; Pred. No. 3.2e-74; Matches 332; Conservative 155; Mismatches 353; Indels 166; Gaps 17;
QY	8	RLACYLEFLKKELEFOLLANKAHSRSSGETPAQPEKTSMEVASVLAQYGEQRAW	67		
DB	7	KLARYLLEDVLDKFKFKHLEDYPPKGCIPLRGQTEKADHVDLATIMDFNGEKA	66		
QY	68	DLALHTWEQMLRSICAQAEQAGHSPFPYSPSEPHLGSPSQPTSTAVLM---PWIHEL	124		
DB	67	AMAVMFAAINRRDLVEKAKR-----DEPKWSDNARVSNPTVICOEDSIEE	114		
QY	125	PAGCTQGSRRVRLQPLDTSGRWRREISASLLYQALPSSPDHSPQSPNAPTSTAVLG	184		
DB	115	WMGLLEYLSRISICKMKDKYKRYKRVRS-RFQCI-----EDNARLGESV--	160		
QY	185	SNGSPPPQSLAREQEAQGTQPLDTSYIYTEREREKSEKGRPPMAAVGTPPOA	244		
DB	161	-----SLNKR-----YTLRLIKEHRSQOEREQELLAIGKTKC	194		
QY	245	HTSLOPHHPWPSPVRESLSTWPMKNEDFNQKFTQLLLQRPHPRSQDPLVKRSPDYV	304		
DB	195	ESPVP-----	200		
QY	305	EENRGLHIBIRLFGPLDTPP-RIVILQGAAGIKSTLARQVKEAWGRGOLYGRFOH	363		
DB	201	-----IKWELLFDPDDESEPHVTVFQGAAGIKTLARKWMLDWSGTLYQDRFDY	253		
QY	364	VYFSCRELAQKYSVLAELIGKDGATAPRIQLSRPERLLFLDGDVDEPGWVLQEPS	423		
DB	254	LFYIHCRESVLTQSRGLIMSCCPDPNPHIKIVKVRSLRILFLMDGDFELQGADEHI	313		
QY	424	SELCLHWSPQPADALLGSLGKTLPEASFLITARTTALQNLIPSLEQARWVEVLGFE	483		
DB	314	GPLCTDQKAERGDILLSSLRKLLPEASLITTPVALEKLOHLLDHPHVEILGFE	373		
QY	484	SRKEYFYRYFTDERQAIRAFRLVKSNEKLWALCLVPWVSWLACTCLMQMKRKEKLTIT	543		
DB	374	AKRKEYFFKYFDEAQAARAFSLIQENEVLFTMCFPLVCWIVCTGLKQKQESGKSLAQ	433		
QY	544	SKTTTTLCHYLAQALQAO-----PLGQLRDLSLAAGSIQKTKLFSPPDLRKHGLD	597		
DB	434	SKTTTAVYVFFLSLLQPRGSGQEHGLCAHLWGLCSLAADGIWNQKILFEESDLRNLHQL	493		
QY	598	GAISTFLKMGILQ-EHPIPLSYSEFIHLCEFFAAMSYYLDEK-GRGKHSNCII----	651		
DB	494	KADVSAFLRMNLFQKEVDCEKFSFIHMTFQFFFAAMYLLLEEKEGRTNVPGRSLKLP	553		
QY	652	-DEKTEIAYG-IHGLFGASTTRFLGLLSDEGEREMENIFHRLSQ--GRNLMQWV----	704		
DB	554	RDVTVLLENYKFEKGYLIFVVRFLPGLVQBERTSVLEKKLSCKISQIRLELLKMWIEVK	613		
QY	705	-PSLQILLPHSLESILHCLYETRNKTLFTQVMAHFEEMCMCVETDMLLVCTFCI-----	758		
DB	614	AKAKLQIQPSQLELYCYLENQBEDFVQRMDFYFPKIEINLSTRMDHNVSSFICENCHR	673		
QY	759	-----KFSRHVKKLOIEGRQHRSTWSPMTVYLFWRVPVTDVYMQ	798		

DB	674	VESLSGLFHNMPKEEBEKEGRHLDWVQCVPSPSSHAACSHGLVNSH-----LTSSPCR 729	
QY	799	ILFSVLKVTNRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLETIRLAGCGLTAEDCKOLA 858	
DB	730	GLFSVLSTSQSLTELDLSGNSLSDGPGMRVLCETLOHPCGNIRRLWLGRCGLSHECCFDIS 789	
QY	859	FGLRANQTLTDLDSFNVLTDAGAKHLQORLQPSCKLQRLQVSCGTLTSCCCQDLASVL 918	
DB	790	LVLSSNQKLVELDLSNALGDFGIRLLCVGLKHLNLLKMLVSSCCLTSACCCQDLASVL 849	
QY	919	SASPSLKELDIQQNLLDDVGVRLLCCEGRHPCAKLIRLGLDQTLLS 964	
DB	850	STSHSLTRYLVGENALGDSVAILCEKAKNPQCMLQKLGVLNSGLT 895	
RESULT 4			
CIS1 MOUSE			
ID	CIS1 MOUSE	STANDARD;	PRT; 1033 AA.
AC	Q8R4B8;		
DT	15-SEP-2003 (Rel. 42, Created)		
DT	15-SEP-2003 (Rel. 42, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing		
DE	APAF1-like protein 1) (Mast cell maturation inducible protein 1).		
GN	CIA51 OR PYPAP1 OR MMIG1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALE/CJ;		
RA	Kiuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,		
RA	Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,		
RA	Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;		
RT	"Identification of inducible genes during in vitro maturation of mouse		
RT	bone marrow-derived mast cells to connective tissue-type mast cells."		
RT	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
RL	!- FUNCTION: May function as a potential inducer of apoptosis.		
CC	Interacts selectively with apoptosis-associated specklike protein		
CC	containing a CARD domain (ASC). This complex may function as an		
CC	upstream activator of NF-kappaB signaling (By similarity).		
CC	!- SIMILARITY: Contains 1 DAPIN domain.		
CC	!- SIMILARITY: Contains 1 NACHT domain.		
CC	!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
DR	EMBL; AF486632; AAL90874.1; --		
DR	MGI; MGI:2653833; Ciasl		
DR	InterPro; IPR001611; LRR.		
DR	InterPro; IPR007091; LRR_RNinh.		
DR	InterPro; IPR003590; LRR_RNinh_sub.		
DR	InterPro; IPR007111; NACHT_NTPase.		
DR	InterPro; IPR004020; PAAD_DAPIN_dom.		
DR	Pfam; PF02758; PAAD_DAPIN_1.		
DR	SMART; SM00368; LRR_R1; 1.		
DR	PROSITE; PS50824; DAPIN; 1.		
DR	PROSITE; PS50837; NACHT; 1.		
KW	Apoptosis; Repeat; Leucine-rich repeat.		
FT	DOMAIN 1 91		
FT	DOMAIN 216 532		
FT	REPEAT 737 760		
FT	REPEAT 794 817		
FT	REPEAT 851 874		
FT	REPEAT 880 903		
FT	REPEAT LRR 1.		
FT	REPEAT LRR 2.		
FT	REPEAT LRR 3.		
FT	REPEAT LRR 4.		

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- ALTERNATIVE PRODUCTS;

CC Event=alternative splicing; Named isoforms=3;

CC Comment=Experimental confirmation may be lacking for some isoforms;

CC Name=1;

CC IsoId=Q96MN2-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q96MN2-2; Sequence=VSP_003917;

CC Name=3;

CC IsoId=Q96MN2-3; Sequence=VSP_003916;

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.

CC -!- CAUTION: A stop codon in Ref.4 was read through in position 41 to extend the sequence.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF42488; AAL35293.1; -;

DR EMBL; AF072792; AAL68396.1; -;

DR EMBL; AF479747; AAL87104.1; -;

DR EMBL; AF482706; AAL88672.1; ALT_INIT.

DR EMBL; AK056688; BAB71254.1; -;

DR EMBL; BC016443; AAL16443.1; -;

DR InterPro; IPR007091; LRR_RNinh.

DR InterPro; IPR007111; NACHT_NTPase.

DR InterPro; IPR004020; PAAD_DAPIN_dom.

DR Pfam; PF02758; PAAD_DAPIN; 1.

DR PROSITE; PS0824; DAPIN; 1.

DR PROSITE; PS0837; NACHT; 1.

DR Repeat; Leucine-rich repeat; Alternative splicing.

FT DOMAIN 1 34 DAPIN.

FT DOMAIN 149 472 NACHT.

FT REPEAT 637 660 LRR 1.

FT REPEAT 698 721 LRR 2.

FT REPEAT 722 745 LRR 3.

FT REPEAT 750 777 LRR 4.

FT REPEAT 806 833 LRR 5.

FT REPEAT 863 886 LRR 6.

FT REPEAT 920 943 LRR 7.

FT REPEAT 949 972 LRR 8.

FT VARSP LIC 1 93

MAASFDGLMWYLEELKKEERKFKHLKQMTLOLELQ

IPWTEVKKASRELANLLIKHVEEQAWNITRIQKMDRK

DLCMKVMBERT -> MQECUTLWVFSPLALTD (in isoform 3).

/FTid=VSP_003916.

Missing (in isoform 2).

/FTid=VSP_003917.

L -> P (IN REF. 4).

Y -> H (IN REF. 4).

I -> T (IN REF. 4).

W -> R (IN REF. 4).

I -> V (IN REF. 4).

Y -> C (IN REF. 4).

L -> P (IN REF. 4).

SEQUENCE 994 AA; 113444 MW; 1CEFC68B18398C9 CRC64;

Query Match 13.8%; Score 1037; DB 1; Length 994;

Best Local Similarity 25.5%; Pred. No. 1.1e-58;

Matches 303; Conservative 168; Mismatches 313; Indels 402; Gaps 25;

9 LACYLFLKKEELKEPQLLANKAHRRSS-----GETP-AQPEKTSNGEVASYLVQVG 62

11 LMWYLEELKKEERFKP-----KEHLKQMTLOLELQIPWTEVKKASRELANLLIKHVE 64

63 EORAWDLALHTWEOGMLRSLCAQA-QEGAGHSPSPFPSPSEPHLGLSPSQPTSTAVLPMWI 121

65 EQAWNITLRIFFQKMDRKDLCKVMWERTGVTKY----- 99

122 HELPAGCTQGSERRVRLQPLDTSGRWRREISASLLYQALPSPDPHESPSQSPNAPTSTA 181

100 -----QAHAKQKFSRL----- 110

182 VLGSWSPQPSLAPREQEAPQTQWPLDTSYIYTEIREREREKSEKGRPPWAAVVGTP 241

111 -----WSSKS-----VTEI----- 119

242 POAHTSLOPHHPWEPVSRESLCTWPKWKNEDFNOKFTQLLLQLRPHPRSDPLVKRWP 301

120 -----HLYPEEVKQSEC----- 132

302 DYVENRGHLEIRDFGGLDTPQPRVILIOGAAGIOKSTLARQVKAWEAGQYIGDRF 361

133 DHLDR-----LPAPKETGQPTVIIQQPGQIGKTTLLMKLMMAWSDNKIFRDRF 182

362 QHVFVFCRELAQSKVSVLAELIGKGTATAPIRQILSRPERLLFILDGVDPCWQLOE 421

183 LYTFYFCRELRUPPTSLADLISREWPDPAPIEIVSQPERLLFVDSFEELGGUNE 242

422 PSSELCHWSQPQADALLGSLGKTIIPASFLITARTTALQNLIPSLQEARWVEVLGF 481

243 PSDLCCGLMEKRPVQVLLSLLRKMLPEASLLIAIKPVCPELKDQVTISEIYQPRGF 302

482 SESSRKEYFYPTDERQAIAPRLVKNKELMALCLVPVWSWLACTCLMQMKRKEKUT 541

303 NESDRLVYFCGFFKDPKRAEAFNLVRESEQLFSICQIPLLCWIILCTSLKQEMQKGDIA 362

542 LTSKTTTTLCLHYLAQAQPL-CP-----QLRDLCSLAAGIWKTKTLPSPDRLKH 594

363 LTCSTTSVSYSFVNLTPEGAEGPTQTOHLKALCSLAAGMWTTFECCEDDLARN 422

595 GLDGAIISTFLKMGILQEH-PIPLSYSIHLFCFOEFFAAMSIVYVLEDEKGRGHNS- - - -C 649

423 GVDADIALLGTLKLYGERESSYVFLAVCIQEFCAALFYLL- - -KSHLDHPHAPVRC 479

650 IDL-----EKTLEYGTHGLFGASTTTRFLGLSLDSEGEREMENIFHCRLSQ- - -GRNLMQ 702

480 VQELLVANFEKARRAHI- - -FLGC- - -FLATGLLNKKEQEKLDAPFGFQSLQEIQQKHQ 533

703 WVPSLQLLLQPH- - -SLESHLCLVETENKTFELQVMAHFEEMGCMVETDMLVCTFCI 758

534 CLKSLGERGNPQGVDSLAIFYCLFEMQDPAPVQAVNLLQENAFHIDNVLVVSAYCL 593

759 KF-----SRHVKKQLQIEGRQHS 777

594 KYCSLRKLCSVQNVFKKEDEHSDTSYSLICWHHICSVLTTSCHLRELQVQDSTLSES 653

778 T-----WSPTMVVLFWRVPVTDAYWQILFSVLKVR- - - - 808

654 TFWTCNQLRHPSCRLOKLGINNVVSFGQSLLFEVLFPQDLKYLSTLTKLSRDDIRS 713

809 -----NLKE----- 812

714 LCDALNYPAGNVKELALVNCHLSPIDCEVLAGLTLNNKLTLYNVNSCNQDPTGVPLCEA 773

813 -----LDLSGNSLSHSAVSKSLCKTLRR 834

774 LCSPTVLVYLMALFCHLSEOCCEYISEMLLRNKSVRVLDLSANVLKDEGLKTLCEALKH 833

835 PRCLLETURLACGGLTAEDCKDLAPGLRANQTLFELDSLNVNLTADAKHLCQRURQSPC 894

834 PDCCLDSLCLVKCFITTAAGCEDLASALISNQNLKIQICNEIGDVGVQLLCRALTHDC 893

895 KLORQLVSCGLTSDCCODLASVLSASPSLKLDELQNNLDDVGVRLICEGRHHPACKLI 954

894 RLEILGLEBEGTSTCCCKDLASVLTCTKLOOLNLTNLTLDHTGVVVVJCEALRHPECALQ 953

955 RLGLDQTTLSDMRQELRALEQEKPLLIFSRRRKPSVMTPTGLDT 1000

DB 954 VLGRKTDDEETQALLTAEERNPLTI-----TDDCDT 988

RESULT 6

MATE_HUMAN
ID MATE_HUMAN STANDARD; PRT; 1200 AA.
AC P59047;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Water protein homolog.
GN MATE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=21922687; PubMed=11925379;
RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
RT "A human homologue of mouse Mater, a maternal effect gene essential
RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
CC -!- TISSUE SPECIFICITY: Oocyte-specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and that statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY054986; AAL15549.1; --
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR RNinh.
DR InterPro; IPR003590; LRR RNinh sub.
DR InterPro; IPR007111; NACHT NTPase.
DR InterPro; IPR004020; PAAD_DAPIN dom.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00368; LRR_R1; 11.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 57 148
FT DAPIN.
FT DOMAIN 280 602
FT NACHT.
FT REPEAT 704 727
FT LRR 1.
FT REPEAT 730 753
FT LRR 2.
FT REPEAT 780 803
FT LRR 3.
FT REPEAT 809 832
FT LRR 4.
FT REPEAT 836 863
FT LRR 5.
FT REPEAT 865 892
FT LRR 6.
FT REPEAT 893 916
FT LRR 7.
FT REPEAT 950 973
FT LRR 8.
FT REPEAT 979 1002
FT LRR 9.
FT REPEAT 1007 1034
FT LRR 10.
FT REPEAT 1036 1059
FT LRR 11.
FT REPEAT 1064 1092
FT LRR 12.
FT REPEAT 1121 1142
FT LRR 13.
FT NP_BIND 286 293
FT ATP (POTENTIAL).
SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A771B28FA CRC64;

Query Match 13.0%; Score 983; DB 1; Length 1200;
Best Local Similarity 25.4%; Pred. No. 4.3e-55;

Matches 329; Conservative 176; Mismatches 411; Indels 380; Gaps 39;
QY 5 AWRLACYLEFLKKEBELKEFOLLANKRAHSSSSGETPAQPEKTSQMEVASYLVAQ-YGE 63
DB SYGLQWCLYE-LDKEEFQTFKELLKKKS-SESTTCSIPOFBIENANVECLALLHEYIGA 119
QY 64 QRAWDLAHTEQWGLRSLCAQAEAGCHSFFYPSPSEPHLGSQSTSTAVLMPWHE 123
DB SLAWATSIISIFENMNLRTLSKARDMKR-----HSPEDPEATMTDQGPS-----KE 166
QY 124 LPAGCTGSSRRVLRQLPDTSGRRWREISASILYQALPSSPDHESPSQESNPATSTAVL 183
DB KPGISQAVQ-----DSATAETKEGEISQA----- 193
QY 184 GSWGSPQPSLAPREQEAGTQWPLDETSGIYIETIREREREKSEKGRPPMAAVGTTPQ 243
DB MEQEGA-----TAAETEEOEISQAMQEGATAETEEQ 226
QY 244 AHTSLQPHHPWPSVRESLCTWPWKNEFNQKFTQLLLQRPDRSQDPLVKES----- 299
DB GHGG-----DTWDYKSHVMT-KF-----AEEEDVRRSFNT 256
QY 300 ---WPDYVEENRGHLIEIRDLFGPGLDTQ---EPRIVLOGAAGIGKSTLARQVKEAWG 352
DB AADWP-----EMQTLAG-AFSDRWGFRPTVVLHGKSGIGKSALARRIVLCWA 304
QY 353 RGQLYGDRFQHFVYFSCRELAQSKVVSIAELTGKGTATPAPIRQILSRPERLLILGV 412
DB OGGLYQGMFYSVFFLPVREMQRKKESSVTEFISREMPDSQAPVTEIMSRPELLIIDGF 364
QY 413 DEPGVLOEPSSELCUHWSPQPADALLGSLGKLTILPEASFLITARTALONLPSLEQ 472
DB DLGSLVANN-DTKLCKDWAQKQPPPTLIRSLIRKLVLLPESFLITVTRVDVGTEKLSEVVS 423
QY 473 ARWTEVLGFSSESSKEYFYRYFTDERQAIRAFRLVKSNNKELWALCVPMVSWLACTCLMQ 532
DB PRYLVIRGIGSEQRILLERIGEHQTKQGLURAIINRELLDQCVFVAVGLICVALQ 483
QY 533 QMKRKEKLTLSKTTTTLCLHYLAQALQAQPLG-----PQLRDLCSLAAEGIWQ 581
DB QDVVCEVAPFNQTLTG--LHAFAFHQLTFRGVVRCNLLEERVVLKRCFCEVGVVN 541
QY 582 KKTLPSPDLRKHGDLGAIISTFLKQILQEHPIPLS-----YSFIHLCFQFFFAAMSV 636
DB RKSVFDGDDLMVQGLGESELRALFHMNII---LPDSHCEERYTFFHLSDQFCALYYV 597
QY 637 LEDEKRGCKHNCIIDLKTLTA-----YGIHGLFGASTTRFLLGLLSDGEREMENI 689
DB LEGLE--IEPALCPLYVEKTRSMELKQAGFHIHSLW---MKRFLGLVSEVDRRPLEVL 652
QY 690 PHCRLSQG--RNLMQWVPSLQLLQPH-----SLESILHCLYETRNKFTLQVMAHFEM 741
DB LCCPVLGVKQKLLHWVSLG--QQPNATTQDITLDAFHCLFETQDKFVRLALNSFQEV 710
QY 742 GNCVETDMELVCTFCIKFSRHVKQLQ-LIEG---RQHRSTWSPMTVVLFRVWVPTDAYW 797
DB WLPINQNLDLTIASSFCLOHCPYLKIRVDVKGFPRDESAEACPVVPLMMRDKLTIEQW 770
QY 798 -----QILF 801
DB EDFCSMLGTHPHLRQLDLGSSILTERAMKTLCAKLRHPTCKTQTLMFRNAQITPGVQHLW 830
QY 802 SVLKVTRNLKELDLSGNSLSHSAVSKLCKTRRRPCLLETTLRAGCGLT----- 850
DB RIVMANRNLRSNLGGLTHKEEDVRNACEALKHPKCLLESRLDCCGLTHACYLKISQIL 890
QY 851 -----AED----- 853
DB TTSPLSKLSLAGNKVTDQGVTPLSDALRVSQALQKLILEDGCGITATGCGQSILASALVN 950
QY 854 -----CKDLAFGLRANQTL 867
DB RSLTHLCLSNNSLGNVLLCRSMRLPHCSLQRLMLNCHLDTAGCGSLALMAGNSWL 1010

```

QY 868 TELDLSFNVLTDAGAKHLQRLQPSCKLQRLQVSCGLTSDCCODLASVLGASPSLKL 927
Db 1011 THLSLSMNPVENGKVLCEVWREPSCHQDLIELVKCHLTAACCSLSVCVSRHLSKL 1070
QY 928 DLQONLNDVGVRLCEGLRHPACKLIRLGLDQTTLSDEMROELRALEOKPQLLIFSR 987
Db 1071 DLTDNALGDDGVAALCEGLKQKNSVLRLGLKACGLTSDCCALS- 1123
QY 988 KFSVWTPTEGLDTGEMNSSTSLKQRLGSERAASHVAQANLKLVDVK-IPF--IAETA 1044
Db 1124 LTSL-----NLVQNNSFSPKMMKLSAFAC---PTSNLQIIGLWKQVYQVQIRKLL 1171
QY 1045 ESSSPVWVVELLCVSPASQGDHLTKPLGTDDDFW 1080
Db 1172 EE-----VQLL-KPRVVIDGSWHS--FDEDDRYW 1197

RESULT 7
NAL2 HUMAN
ID NAL2 HUMAN STANDARD; PRT; 1062 AA.
AC Q9NX02; Q9BVN5; Q9H6G6; Q9HAV9; Q9NWK3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NACTH-, LRR- and PYD-containing protein 2 (Nucleotide-binding site
protein 1) (PYRIN-containing APAF1-like protein 2).
GN NALF2 OR NES1 OR PYPAF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21169419; PubMed=11270363;
RA Bertin J., DiStefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
proteins.";
RL Cell Death Differ. 7:1273-1274 (2000).
[2]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=21148093; PubMed=11250163;
RA Martinon F., Hofmann K., Tschopp J.;
RT "The pyrin domain: a possible member of the death domain-fold family
implicated in apoptosis and inflammation.";
RL Curr. Biol. 11:R118-R120 (2001).
[3]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
activation of NF-kappa B and caspase-1-dependent cytokine
processing.";
RL J. Biol. Chem. 277:29874-29880 (2002).
[4]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Colon, and kidney epithelium;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
Nakamura Y., Isigai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung, and placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Muller R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Klausner R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Boesk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield A.Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: May be implicated in apoptosis (By similarity).
CC -!- COFACTOR: Binds ATP (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NX02-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NX02-2; Sequence=VSP_005522;
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACTH domain.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF298547; AAG15253.1; ALT_INIT.
DR EMBL; AF310106; AAG30289.1; -
DR EMBL; AAL64764; AAL69962.1; -
DR EMBL; AK000517; BAA91223.1; -
DR EMBL; AK025952; BAB15293.1; -
DR EMBL; AK000784; BAA91377.1; ALT_INIT.
DR EMBL; BC003592; AAH03592.1; -
DR EMBL; BC001039; AAH01039.1; -
DR HSSP; P10775; 2BNH.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACTH_NTFase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACTH; 1.
KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
FT Alternative splicing.
FT DOMAIN 1 94 DAPIN.
FT DOMAIN 207 526 NACTH.
FT REPEAT 467 491 LRR 1.
FT REPEAT 622 645 LRR 2.
FT REPEAT 754 777 LRR 3.
FT REPEAT 810 832 LRR 4.
FT REPEAT 839 862 LRR 5.
FT REPEAT 867 890 LRR 6.
FT REPEAT 924 947 LRR 7.
FT REPEAT 981 1005 LRR 8.
FT REPEAT 1010 1033 LRR 9.
FT REPEAT 1010 1033 ATP (POTENTIAL).
FT DOMAIN 213 520 POLY-GLU.
FT DOMAIN 518 523 Missing (in isoform 2).
FT VARSPLIC 133 154 /FTId=VSP_005522.
FT CONFLICT 1 1 M -> V (IN REF. 1).
FT CONFLICT 35 35 L -> P (IN REF. 1).
FT CONFLICT 304 304 I -> S (IN REF. 4; BAB15293).
FT CONFLICT 364 364 R -> K (IN REF. 1).
FT CONFLICT 980 980 MISSING (IN REF. 1).
FT CONFLICT 1052 1052 A -> E (IN REF. 5).
FT CONFLICT 1062 AA; 120514 MW; 4DBB0F6E9C2BC8A7 CRC64;
SQ SEQUENCE

```


Query Match		12.6%;	Score 948;	DB 1;	Length 1062;
Best Local Similarity		27.8%;	Pred. No. 6.4e-53;		
Matches		302;	Conservative	167;	Mismatches 389; Indels 228; Gaps 32;
QY	148	WREISALLYQALPSSPDHESQSPNAPTAVLGSWGSPPQSLAPREQEAFTQWP	207		
DB	67	WVEMASLQVFEK-----HRMDSEAKDEVREAAALKSFNKKRKLUSLGIITRKRP-----P	117		
QY	208	LDSETSIIYTEIRERSEKSEKRPWAAVAVGTPPOAHTSLOPH-----HHPWEPSPVRES	262		
DB	118	LD-----VDEMLERKTEQAFTETKGNVICLGEVFKGKPKDKDNCRYLTKTKFRE-	170		
QY	263	LCSTWPKNEDFN---OKTQTLLOLQRPHPRSODPLVKRSWPDYVEENRGHILIEIRDPLFG	319		
DB	171	MNKSPPGDSKEVQVMAERYKMLIPFNP-----RVLPQ	203		
QY	320	PGLDTQEPRIIVILQAGAGIKSTLARQVKEAWGRGOLYGDRLFQHVFPYFSCRELAQSKVVS	379		
DB	204	PFSYT-----VVLYPAGLGTTLAQKMLMDWAEDNLI-HKPKYAFYLSRSLSLGPGCS	257		
QY	380	LAELIKGCTATPAPTRQILSRPERLFIIDGVDE----PGWVLOEPPSELCLHWSQOP	435		
DB	258	FIELVFRDNPLODDIPHILAQARKILFVIDGFDELGAAPGALIE-----DICDWEKKP	313		
QY	436	ADALLGSLGKTLIPASFLITARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYFT	495		
DB	314	VPVLLGSLNRYMLPKAALLVTRPRALRDLRLAIEPIYIRVEGFLEDDRDAYFLRHFG	373		
QY	496	DERQAIRFLVSKNELMALCLVPVWSWLAICTCLMQMKREKLTLSKTTTTLCLHVL	555		
DB	374	DEDQAMRAPELRSNAALFQASAPAVCIWTTTLQMEKGEDPVPTCLTRTGLFLRFL	433		
QY	556	AQAL-QAQLPGPOLRLCSLAAGIWKTLFSPDRLRKHGLDGLATISFLKMWILOEHP	614		
DB	434	CSRFOGAQRGAIRLTSLLAQGLWAQTSVLHREDLERLQVQESDLRFLDGLILRQDR	493		
QY	615	IPLS-YSFHLCPOEFFAAMSYVL--EDBKGRGKSHNCIIDLKTELEAYGIGHLFGCAST	671		
DB	494	VSKGCYSFIHLSFQQLFTALFYLTXEEDRDGHTWDIGVQKLLS--GVERLRNPDLI	551		
QY	672	R---FLLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSL-QLLQ-----PHSL----	716		
DB	552	QAGYVSFGLANERAKRELEATFCRMS-----PDIKQELLRLCDISCKGHSVTVDL	602		
QY	717	-ESLHCLYETRNKFTLTQWMAHEPMGCMVETDMELLVCTCFKFSRHVKLQL-----	769		
DB	603	QELLGCLYSQEBELKEVNAQKREISLHUNA-VDVVPPSFCVKHCRNLQKMSLQVIKEN	661		
QY	770	-----IEGRQ-----HRSTWSPTMV-----	784		
DB	662	LPENVTASBDAEVSQDDQHMLPFWTDLCSIFGSKNDLGLAINDSFLSASLVLCE	721		
QY	785	-----VLFWRVPVTDAYQILSVLKVTRNLKELDLSGNSLSHSAVKSLCKTLR	833		
DB	722	QIASDTCHLQVVVFKNISPADAHNLCLA-LRGKTVTLTLQDND-QDDMPALCEVLR	779		
QY	834	RPCILLETILACGGLTAEDCKDLAFLGRANQTLTDLISFNVLTDAGAKHLQORLQPS	893		
DB	780	HPECNRYLGLVSCSATTOQWADLSALEVNSQSLTCVNLSDNELLDEGAKLLYTLRHPK	839		
QY	894	CKLQRLQVSCGLTSDDCCDLASVLSASPSLKELDLQONNLDVGVRLLCGLRHHPACKL	953		
DB	840	CFIQLRSLNCHUTEANCCKDLAAVLVVSRELTHLCLAKPIGNTGVKFLCEGLRYPECKL	899		
QY	954	IRLGLDQTLSDMRQELRALEQKQLLIFSRKPSVMTPTTEGLDTGMSNSTSLKQ	1013		
DB	900	QTLVLWNCDDITSDGCCDLTKLLQEKSSLLC-----LDLG-----	933		
QY	1014	RIGSERAASHVAQANIKLLDVSKIFPIAEI-----AESSEPVVPVELLCVPSPASQ	1066		
DB	934	-----LNHIGVGMKFLCEALRKPLCNLCILMWGCSIPPFSCDCLCSALSCNOSLVT	986		

QY	1067	-DLHTKPLGTD-----DDFWGPTGVPVATEVVDKEKNLY---R	1099		
DB	987	LDLQGNPLGSGVGMVLFETLTCSSGTLRLKIDDFNDELNKLLEIEEKNPQLIIDTE	1046		
QY	1100	VHFPVA 1105			
DB	1047	KHPWA 1052			
RESULT 8					
PYA3_HUMAN		STANDARD;	PRT;	980 AA.	
ID	Q8WX34;				
AC	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	PYRIN-containing APAF1-like protein 3.				
GN	PYPAF3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22162427; PubMed=12019269;				
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,				
RA	Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.,				
RT	"PYPAF7, a novel PYRIN-containing APAF1-like protein that regulates				
RT	activation of NF-kappa B and caspase-1-dependent cytokine				
RT	processing.";				
RL	J. Biol. Chem. 277:29874-29880(2002).				
CC	- SIMILARITY: Contains 1 DAPIN domain.				
CC	- SIMILARITY: Contains 1 NACHT domain.				
CC	- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; AF464765; AAL69963.1; -				
DR	InterPro; IPR007091; LRR RNinh.				
DR	InterPro; IPR007111; NACHT NTPase.				
DR	InterPro; IPR004020; PAAD DAPIN_dom.				
DR	Pfam; PF02758; PAAD DAPIN; 1.				
DR	PROSITE; PS50824; DAPIN; 1.				
DR	PROSITE; PS50837; NACHT; 1.				
KW	ATP-binding; Leucine-rich repeat; Repeat.				
FT	DOMAIN 1 93 DAPIN.				
FT	DOMAIN 172 491 NACHT.				
FT	REPEAT 614 638 LRR 1.				
FT	REPEAT 674 697 LRR 2.				
FT	REPEAT 760 784 LRR 3.				
FT	REPEAT 788 810 LRR 4.				
FT	REPEAT 817 840 LRR 5.				
FT	REPEAT 845 868 LRR 6.				
FT	REPEAT 874 897 LRR 7.				
FT	REPEAT 902 928 LRR 8.				
FT	REPEAT 933 957 LRR 9.				
FT	NP_BIND 178 185 ATP (POTENTIAL).				
SQ	SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;				
Query Match 12.1%; Score 908.5; DB 1; Length 980;					
Best Local Similarity 30.2%; Pred. No. 2e-50;					
Matches 261; Conservative 128; Mismatches 284; Indels 191; Gaps 21;					
QY	294	PLVKRSGWPDYVEEN-----RGLHIEIRD-----LFGPGLDTQEP	332		
DB	117	PGEKSGWRNMEKQSLVKNWNTFQGDIDFNHDDVTLRNQRFIFPLNPRPKLPTTYVL	176		

333 QGAAGIGKSTLARKVKEAGRGOLYGRFQHVFFYSCRELAOSKVSLAELIGKDGATP 392
 177 HGPAGVGKTTLAKKMLDTCNL-SPTLRYPYLSCKELSRMGPCSAELISKDWPELQ 235
 393 APIRQILSRPERLLFLDGVDE----PGWVLQBPSSSELCHWSQOPADALLGSLGKTY 448
 236 DDIPSILAQAQRTLVFVVDGLDEKVPFGALIQ----DICGDWEKKKVPVLLGSLKRXK 291
 449 LPASFLITARTALONLPSLEBOARWVEVLFSSSRKEYFYRYFTDERQAIRAFRLVK 508
 292 LPRAALLVTRPRALRDQLLQAQPIYVVRGEFLEEDRRAYFLRHFGEDDQWRAPFELMR 351
 509 SNKELWALCLVPVWSVLACTCLMQMKRKEKLTLSKTTTTLCHYLAQAL-QAQLPGSQ 567
 352 SNAALFQLGSAPAVCIVCTTLKQMEKGEDPVPTCLTRTGLFLRLCSRRFQGAQLRCA 411
 568 LRDLCSLAEGIWKKTLSFPDDLKHGIDGAIISTFLKMGILQHPILS-YSFIHLCF 626
 412 LRTLSLAAGLQAQWQSVFREDLERLGVQESDLRLFLDGLDILQDRVSKGCYSFIHLSF 471
 627 QBFPAAMSVYLEDEKG--RGKHSNCIIDLKTELEAVGIHGLFGASTR---FLGLLSDE 681
 472 QQLTALFVALEKEGEDRDGHAWDIDGVOKLLS--GEBRLKNPDLIQVGHFLFLGANEK 529
 682 GEREMENIFHRLSQGRNLMQWVPSL-QLLQ-----PSHL-----ESLHCLYETRN 727
 530 RAKELEATFCRMS-----PDIKQELLQCAHLHANKPLSVTDLKEVLGCLYESQE 580
 728 KFTLTQWMAHFEBMGCVETDEMLLVCTTCIKFSRHVKLQL-----769
 581 EELAKVVVAPFKBISIHILTNTSEVMHCSFSLKHCQDLQKLSLQVAKGVFLENYMDFELDI 640
 770 -----LEGRQH-----775
 641 EPERCTYLTIPNWARODLSRLWTFCSLFSSNSMLKFLVQKQSPFLSDSSVRILCDHVT 700
 776 RST-----WSPTMVVL-----786
 701 RSTCHLQKVEIKNVPTDVAIROFLAFIGKKTTLTLTLAGHIEWERTWMLCDDLNRNHK 760
 787 --FRWVPV-----TDAYWOILSVLVKTRNLKELDLSGNSLSHSAVSKLCTLRPRCLL 839
 761 CNLQYLRLGCHCATPQWAEFFVYLVKANQSKHLRLSANVLDEGAMLLYKMTTRPKHPL 820
 840 ETURLAGCGLTAECDKDLAFGRANOTLTLELDLSFNVLTDAKAKHLQRLRPSCKLQRL 899
 821 QMLSENCLRTASCKDLAAVLVSKGLTHLCLAKNPIDGTGVKFLCEGLSPDCKLQTL 880
 900 QLVSCGLTSDCCDLASVLSAPSLKELDLQNNLDDVGVRLICEGLRHPACKLIRLGLD 959
 881 VLQOCSITKLGCRYLSEALQEAACSLTNLDLSINQIAR-GWLILQALENPNCLNKLRLK 939
 960 QTTLSDEMROELRALEQKPKLLI 983
 940 TYETNLEIKKLEEVKEKPKLTI 963

RESULT 9

PYAS_HUMAN STANDARD; PRT; 892 AA.
 ID PYAS_HUMAN
 AC P59044;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 5.
 GN PYPAP5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22162427; PubMed=12019269;

RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.,
 RT "PYPAP7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing."; J. Biol. Chem. 277:29874-29880(2002).
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=22275822; PubMed=12387869;
 RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
 RA Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
 RA Distefano P.S., Bertin J.;
 RT "Functional screening of five PYPAP family members identifies PYPAP5
 RT as a novel regulator of NF-kappaB and caspase-1."; FEBS Lett. 530:73-78(2002).
 RL FEBS Lett. 530:73-78(2002).
 CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
 CC activation of NF-kappa-B.
 CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
 CC much lower levels in T-cells.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; AF479748; AAL87105.1; -;
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_R1; 3.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.
 KW DOMAIN 1 103
 FT DOMAIN 1 103
 FT DOMAIN 196 513 NACHT.
 FT DOMAIN 504 614 POLY-GLU.
 FT REPEAT 462 487 LRR 1.
 FT REPEAT 727 747 LRR 2.
 FT REPEAT 755 778 LRR 3.
 FT REPEAT 811 834 LRR 4.
 FT REPEAT 845 868 LRR 5.
 FT NP_BIND 202 209 ATP (POTENTIAL).
 SQ SEQUENCE 892 AA; 98733 MW; 4AA8D1FC766DDE9D CRC64;
 Query Match 12.0%; Score 901.5; DB 1; Length 892;
 Best Local Similarity 27.8%; Pred. No. 4.8e-50;
 Matches 295; Conservative 143; Mismatches 328; Indels 295; Gaps 33;
 QY 9 LACYLEFLKKEELKEFQALLANKAHSRSSOET-P-AQPEKTSQMEVASYLVQYGEORA 66
 DB 22 LLALEELSQELKFRF---HKLRDVGPDGRSIPWGRLEERADAVDLAEQLAQFYGPSPA 77
 QY 67 WDLALHTWEQGLRSLCAQAOE-----GAGHSFPFVSPSEPHLGSQSTSTAVLMPW 120
 DB 78 LEVARKTLKRADARDVAQQLQERLQRLGLSGTLLSVSEYK-----119
 QY 121 IHELPAGTCQSERRVQLPDTSGRRWRREISASLLYQALPSSPDHESPSQESPNAPTST 180
 DB 120 -----KKYREHVQLQHARVKNARSVKIKRFTKLLIAP-----155
 QY 181 AVLGSWSPQPSLAPRQEAQPTQWPLDETSGIYYTIREREREKSEKGRPPAAVVGVT 240
 DB 156 -----SAAPEALGPAAEPEFG-----RARRSD-----178

241 PQAHTSLQPHHPWSPVRESLCTWPKWEDFNQKFTQLLLQRPSPQDPLVGRSW 300
 179 ---THT---ENLPRR-----188
 301 PDYVENRGHLIEIRDLFGGLDQEPRIIVIOGAAGIKSTLQVKEAWGRGOLYGD 360
 189 ---DEGR-----RPLTVIOGPGAGIKGTAAKILVDMAAGLYQOQ 228
 361 FQHVFFVSCRELAQSK-VVSLAELIKGDTATPAPIRQILSPERLLFILDGVDE-PGWV 418
 229 VDFAFMPGCELLERPGTSLADLILDCPDGAPVQMLAQORLLFILDGADLP--A 286
 419 LQEPSELCHWSQPOPADALLGLSLGKLTILPEASELITARTALONLIPSLQEARWVEV 478
 287 LGSPEAAPCTDPEAASGARVLGGLSKALLPTALLLVTTAAAPGLRGLCSQCAEV 346
 479 LGESESRKEYFRRYFTDROAIRAPRLVYKSKELMALCLVPMVSWLACTCLMOQKRKE 538
 347 RGFSDKDKKKYFKFRDERRAEARVFKENETLFALCFVPVCMVIVCTVLRQQLLGR 406
 539 KLTLSKTTTTLCHVLAQALQAPL--GPO---LRDLCSLAAEGIMQKTLFSPDILR 592
 407 DLSRTSKTTTIVLLFITSVLSAPVADGPRLOGLDLRLCLAREGLVGRRAQFAKELE 466
 593 KHGLOGAIIST-FLKM---GILQEHPIPLSYSFHLCQEFPAAMSVYLED-----E 640
 467 QLELRGSKVQTLFLSKKELPGVLETE--VTYQFIDQSQEFALSYLLEDGGVPRTA 523
 641 KERG-----KHSCNIIDLEKLEAYGIHGLFGASTTRPLGLLSDGEREMENIFHC 692
 524 GGVGTLLRGAOPHSHVL-----TTRFLGLLSAERMDIERHFQC 565
 693 RLQO--GRNLMOWV-----PSLQLLQP-----HSLSLHCL 722
 566 MVSERVKQALRWQOQGGCPQVAVEVEGAKLEDTEPEEBEGBEPNPLELLYCL 625
 723 YETRNKTLFTQVMAHFEEMGM-----CVETDBELLVCTCFIKFRSVKVKLQIEGRQRS 777
 626 YETQEDAFVRQALCRFPALQVRFC--RMDVAVLSYVCRCPAGQALRLISCR---- 678
 778 TWSPTWVLFVRVPTVDAYWQILFVLKVTNLI---KELDLSGNSLSHSAVSKLCTLR 834
 679 -----LVAQEKKKSKLGLKGLASGLSGSSQGTTRQLPASLH 717
 835 -----PRCILETLRAGCGLTAECDKDLAFGLRANQTLFELDLSEFNVLTDAGAKHL 886
 718 PLFQAMTDPCLHSLUSLTHCKLPDAVCRDLSEALRAAPALTELGLHLNRLSEAGLRMUS 777
 887 QRLRQPSCKLIQ--RLQL--VSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGRILL 942
 778 EGLAWPQCRVQTVRVQLPDPQRLG-----QYLVGMLRQSPALTTLDLSCQLPAPMVTYL 832
 943 CEGLRHPACKLIRLGLDQTLTSEMQRLEALREKRPQLLI 983
 833 CAVLOHQGCGLOTLTSLASVELSEQLQELQAVKRAKPDILVI 873

RESULT 10

PYAS_RAT
 ID PYAS_RAT STANDARD; PRT; 854 AA.
 AC Q63035;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAFI-like protein 5-like (Angiotensin II/vasopressin receptor).
 GN PYAP5 OR AVR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]

RP CONCEPTUAL TRANSLATION OF 1-343.
 RA Hinz U.;
 RL Unpublished observations (FEB-2003).
 RN [2]
 RP SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
 RC TISSUE=Kidney;
 RX MEDLINE=96071640; PubMed=7489366;
 RA Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
 RT "Identification of a novel dual angiotensin II/vasopressin receptor on the basis of molecular recognition theory.";
 RL Nat. Med. 1:1074-1081(1995).
 RN [3]
 RP VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
 RX MEDLINE=21980185; PubMed=11984003;
 RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
 RT "The dual AngII/AVP receptor gene N19S/C163R variant exhibits sodium-induced dysfunction and cosegregates with salt-sensitive hypertension in the Dahl salt-sensitive hypertensive rat model.";
 RL Mol. Med. 8:24-32(2002).
 RN [4]
 RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYAP5.
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
 RT "Identification of mammalian orthologs associates PYAP5 with distinct functional roles.";
 RL FEBS Lett. 538:173-177(2003).
 CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote activation of NF-kappa-B (By similarity). Angiotensin II and vasopressin binding protein. May stimulate CAMP accumulation.
 CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low levels in all tissues tested.
 CC -!- DISEASE: Defects in PYAP5 may be a cause of salt-sensitive hypertension.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
 CC -!- CAUTION: The N-terminus was extended using ESTs and genomic sequences, in analogy to ortholog sequences.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

EMBL; M85183; AAA03623.1; ALT_INIT.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 103 DAPIN.
 FT DOMAIN 168 484 NACHT.
 FT DOMAIN 564 604 ASP/GLU-RICH.
 FT DOMAIN 655 662 POLY-LYS.
 FT REPEAT 433 458 LRR 1.
 FT REPEAT 610 633 LRR 2.
 FT REPEAT 722 745 LRR 3.
 FT REPEAT 812 836 LRR 4.
 FT NP_BIND 174 181 ATP (POTENTIAL).
 FT VARIANT 492 492 N -> S (IN SALT-SENSITIVE HYPERTENSION).
 FT VARIANT 536 536 C -> R (IN SALT-SENSITIVE HYPERTENSION).
 FT MUTAGEN 766 766 E->K; ABOLISHES ANGIOTENSIN II BINDING.
 SQ SEQUENCE 854 AA; 95292 MW; D7BBD922D7B3734 CRC64;

Query Match 11.9%; Score 895; DB 1; Length 854;
 Best Local Similarity 33.2%; Pred. No. 1.2e-49;
 Matches 261; Conservative 118; Mismatches 298; Indels 110; Gaps 22;

```
QY 275 NOKFTQLLQLRPHRSQDPLVKRSPDYVEENRGHLIEIRDLFGGLTQBP--IVIL 332
DB 114 NKEFTKLLIAPGSGAGEDELLTSGEPERARRSDTHFNRLF-RGNDEGPRPLTVVL 172
QY 333 QGAAGTGTARQVKEANGRGOLYQDRQHFVYFSCRELAQSK-VSLAEILKGDGTAT 391
DB 173 QGPAGTGTAAKLLYDAGGLYHSDQVDFAFPMFCGELLERPGTRSLADLILEQCPDR 232
QY 392 PAPIRILSRPERLLFILOVDV-PCWVLQEPSSSELCHWSQPOPADALLGLSLGKTIPL 450
DB 233 TAPVRILAQPHRLFILOGADEL-PTLAAPEATEPCRPDPFEATSGRLVSLGSLGQELLP 290
QY 451 EASFLITATTALONLIPSLQARWVEVLGFSSESSKEYFYFYFTDERQAIARFLVKSN 510
DB 291 SARLLVTSRNATLGRLOGLRCSPOCHEVRGFSDDKKKYPFFFRDERKAERAYRIVEN 350
QY 511 KELMALCLVPVWSWLAQCTCLMQOMKKEKLTTSKTTTTLCLHYLAQALQ-AQPLGP--- 566
DB 351 ETLVLCFVFFVVCWIVCTVLLQOMELGRDLRSKTTTSVYLLFITSMLKSAGTNGPRVQ 410
QY 567 -QLRDLCSLAAGIWKOKTLFSPDDDLRKHGLDGAIIIST-FLKM---GILQEHPIPLSYS 620
DB 411 GELRMLCRAREGILKHQAFSEKDLERLKLQSGVQVTFWLSKKELPGLVLE---TVVTYQ 467
QY 621 FIHLCFQEPFAAMSYYVLEDEKGRGHSNCIIDLKTEAYGIHGLFG--ASTTRFLGLL 678
DB 468 FIDQSFQFLAALSYLLDAB---GAPGNSAGSVQMLNSDA-GLRGHLATTRFLFGLL 522
QY 679 SDGEREMENIFHC-----RLSQGRNLMQWPSLQLLLP----- 713
DB 523 STERIRDIGNHFGCVVPRVKQ--DTRLRW---QGSQPKVATVGAEKDELKDEABEE 577
QY 714 -----HSLESILCIYETRNKTLFTLQWMAHEEEMGCV-----TDMELLVCTFCI 758
DB 578 EEEEEEEELNFGLELLYCLYETQEDDFVRQALSLPE--MVLERVLRTRMDLEVLSCV 635
QY 759 KFSRHVKLQLI-----EGRHSRTWSPTMVLFRWVPVTDAYQLFSLVKLTRLN 810
DB 636 QCCPDQALRLVSCGLVAKEKKKKSP-----MNL 668
QY 811 KELDLSGNLSHSAVSKLTKLRLRCLLETURLACGLTAEDCKDLAFGLRANOTLTEL 870
DB 669 KGSQSTGKQPASLLRPLCEAMITQCGLSILTLSHCKLPDAVCRDLSEALKVAPSLREL 728
QY 871 DLSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQ 930
DB 729 GLQNRLLTEAGRLLSQGLAWPKCKVQTLRIQMPGL-QEVIHYLVIVLQSQSPVLTLDLS 787
QY 931 QNNLDDVGRLLCEGLRHHPACKLIRLGLDQTTLSDEMROELRALEQEKPOLLIFFSRKPS 990
DB 788 GCOLPTVVEPLCSALKHPKCGLTKLSLTSVELTENPLRELQAVTKLKPDLAIHSLGKT 847
QY 991 VMTPTREG 997
DB 848 HPQPLKG 854
```

RESULT 11

```
MATE_MOUSE
ID MATE_MOUSE STANDARD; PRT; 1111 AA.
AC Q9R1M5; Q9JLR2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Maternal antigen that embryos require (Mater protein) (Ooplasm-
specific protein 1) (Op1).
GN MATER.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chordata; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss; TISSUE=Ovary;
RX MEDLINE=99360614; PubMed=10433232;
RA Tong Z.-B., Nelson L.M.;
RT "A mouse gene encoding an oocyte antigen associated with autoimmune
premenstrual ovarian failure.";
RL Endocrinology 140:3720-3726(1999).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS ASP-7; LEU-42; LYS-87; PRO-256;
RC GLN-977 AND 1003-ASN-ASN-1004.
RX STRAIN=129/SV;
RA MEDLINE=2022768; PubMed=10754103;
RT Tong Z.-B., Nelson L.M., Dean J.;
RL "Mater encodes a maternal protein in mice with a leucine-rich repeat
domain homologous to porcine ribonuclease inhibitor.";
RN Mamm. Genome 11:281-287(2000).
RP [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517328; PubMed=11062459;
RA Tong Z.-B., Gold L., Pfeifer K.E., Dorward H., Lee E., Bondy C.A.,
Dean J., Nelson L.M.;
RT "Mater, a maternal effect gene required for early embryonic
development in mice.";
RL Nat. Genet. 26:267-268(2000).
CC -!- FUNCTION: Necessary for embryonic development beyond the 2-cell
stage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Oocyte-specific.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074018; AAD51762.1; --
DR EMBL; AF143573; AAF64393.1; --
DR EMBL; AF143559; AAF64393.1; JOINED.
DR EMBL; AF143560; AAF64393.1; JOINED.
DR EMBL; AF143561; AAF64393.1; JOINED.
DR EMBL; AF143562; AAF64393.1; JOINED.
DR EMBL; AF143563; AAF64393.1; JOINED.
DR EMBL; AF143564; AAF64393.1; JOINED.
DR EMBL; AF143565; AAF64393.1; JOINED.
DR EMBL; AF143566; AAF64393.1; JOINED.
DR EMBL; AF143567; AAF64393.1; JOINED.
DR EMBL; AF143568; AAF64393.1; JOINED.
DR EMBL; AF143569; AAF64393.1; JOINED.
DR EMBL; AF143570; AAF64393.1; JOINED.
DR EMBL; AF143571; AAF64393.1; JOINED.
DR EMBL; AF143572; AAF64393.1; JOINED.
DR PIR; A59000; A59000.
DR HSP; P10775; 2BNH.
DR MGD; MGI:1345193; Mater.
DR GO; GO:0005829; Cytoplasm; IDA.
DR GO; GO:0009887; P:organogenesis; IMP.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF00560; LRR; 3.
DR PROSITE; PS00837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat; Polymorphism.
FT DOMAIN 1 128 5 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 1 23 1 (INCOMPLETE).
FT REPEAT 24 49 2.
FT REPEAT 50 75 3.
FT REPEAT 76 102 4.
FT REPEAT 103 128 5.
FT DOMAIN 191 513 NACHT.
```

FT REPEAT 640 664 LRR 1.
 FT REPEAT 691 714 LRR 2.
 FT REPEAT 737 762 LRR 3.
 FT REPEAT 776 799 LRR 4.
 FT REPEAT 804 832 LRR 5.
 FT REPEAT 834 856 LRR 6.
 FT REPEAT 861 884 LRR 7.
 FT REPEAT 890 913 LRR 8.
 FT REPEAT 918 941 LRR 9.
 FT REPEAT 950 974 LRR 10.
 FT REPEAT 975 1003 LRR 11.
 FT REPEAT 1005 1027 LRR 12.
 FT REPEAT 1032 1059 LRR 13.
 FT NP_BIND 197 204
 FT VARIANT 7 7
 FT VARIANT 42 42 E -> D (IN STRAIN 129/SV).
 FT VARIANT 87 87 G -> L (IN STRAIN 129/SV).
 FT VARIANT 256 256 R -> L (IN STRAIN 129/SV).
 FT VARIANT 977 977 L -> P (IN STRAIN 129/SV).
 FT VARIANT 1003 1004 H -> Q (IN STRAIN 129/SV).
 FT SEQUENCE 1111 AA; 125502 MW; 4016A5D67A1C01F4 CRC64;
 Query Match 11.7%; Score 882; DB 1: Length 1111;
 Best Local Similarity 26.1%; Pred. No. 1.2e-48;
 Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

QY 162 SSPDHSPOESNPAP-----TSTAVLGSWGSPQPS-----LAPR-----EQEAPGTOWPLD 209
 DB 28 TSPENDSKSIQKQOEQESTSTWG-----PEKDSKAILKARGLEEQKSESTWSPSE 83
 QY 210 ETSGIYYTIRERERKSEKRPAAVGTTPQANTSLQPHHPWEPVSRESLCSTWPM 269
 DB 84 NVSRAILKDSGSBEVQASERK-----NTSPENDSKSIQKQOEQESTWSPSE 133
 QY 270 KNEDFNQKFTQLLLQRPSPQDPLVKSPPYVEENRCHLIE----- 313
 DB 134 KEED-----EVTE-----ADKONGGD-LQDYKAHVIAKFTSDVLHYDSEPMK 175
 QY 314 -IRDLFGPLGDTQPRIVLQAGAGTCKSTARQVKEAWGRGOLYGDGFQHFYFSCREL 372
 DB 176 LLSDAKPKQKTPQHTIILHGRPGVKSALARSIVLGHQAQKLP-QKMSFVIFFSVREI 234
 QY 373 AQSKVSVLAELICKDGTATPAPIROILSRPERLLFDLGDVDEPGWVLPFSPSELCLHWSQ 432
 DB 235 KWTEKSLAQLIAKECPDSWLVTKMSQPERLLFVIDGLDDMSVLQHDHMTLSRDWKD 294
 QY 433 PQADALLGSLICKTLPFASFLITARTALQNLIFSLQEARVVEVLGFSSESRSKRYFYR 492
 DB 295 EQPIYILMYSLLRALLPQSFLITRTNTGLEKLSKMWVSPVLVLEGLSASRRSOLVLE 354
 QY 493 YFTDERQAIAPRLVKSNEKLMALCLVPWVSWLACTCLMOQMKRKEKLTITSKTTT----- 548
 DB 355 NISNESDRIQVHSLIENHQLFQOQAPSVCSLVCEALQKLGKRGRTLPQOGLTGLYA 414
 QY 549 TLLCHYL-----AQALQAQPLGQLRDLCSLAEEGWQKRTLPSPDRLKRGHLDGAIIST 603
 DB 415 TLVHFQTLKRPQSALSQEQITVLGLCMAAEGVWTVESVFYDDLLKNYSLKSEILA 474
 QY 604 FLKMGILQE--HIPILSYSTIHLCPQEFPAAMSVYLEDEKGRKH-----SNCTIDLEK 655
 DB 475 LPHMNIILQVGHNSSEQCYVFSHLSLODFFAALYVLEGLEEMWQHCFCTIENQRISMEVR 534
 QY 656 TLEAYGHGLFGASTTRFLGLLSDGEREMENIFHCRU--SOGRNLMQWVPSLQLLLOP 713
 DB 535 TDDT-----RULGMK--RFLGLMKNKILKTLVLFYFVPTVEQKLOHWV--SLIAQQ 585
 QY 714 -----HSLIESHLIYETRNKTLFTQVMAHFEEMGMCVETDMELLYVCFIKFSRHVKK 766
 DB 586 VNGTSPMDTLDAYCYLFESODEFVGALKRFEVWLLINQKMDLVKSVYCLKHCQNLKA 645
 QY 767 LQL-----IEGRQHRSTWSPMVL-FRWVPVTDAYWQIILFVSLVKVTRNLKELDLNSGUS 821
 DB 646 IRVDIRDLISVDNTLPCPVVTQETQCKPLLMWGNFCSVLGSLRNKELDLGDSILS 705

QY 822 HSAVKS----- 828
 DB 706 QRAMKILCLELRNOSCRIOKLTFFKSAEVVGLKHLWKLKLLFSNQNLKYLNLGNTPMKDDM 765
 QY 829 ---CKTLRPRCLLETIR----- 843
 DB 766 KLACALPKPKSVETLRLDSCELTIIGYEMISTLLITRLLKCLSLAKNRVGVKSMISL 825
 QY 844 -----LAGGLTAEDCKDLAFGLGRANQTLTDLDFSNVLTDAKAKHLCOR 888
 DB 826 GNALSSMCLLOKLLDNCGLTPASCHLLVLSALFSNQNLTHLCLNNSLGTGVQOLCOF 885
 QY 889 LRQPSCKLQ----- 898
 DB 886 LRNPECALQRLTLNHCNIVDDAYGFLAMLANNTKLTLSLTMPVPGDAMKLLCEALKE 945
 QY 899 -----LQLVSCGLTSDCCQDLASVLSASPSLKELDLQONNLDVGVRLLCGLRHPAC 951
 DB 946 PTCYLQELVDCQLTQCCEDLACMITTTTKLKSLLDGLNNAALGDKGVITLCEGLKQSS 1005
 QY 952 KLIRGLQDTLLSDMROELRALEQEKPOL 981
 DB 1006 SLRRLGLGACKLTSNCCALSALSCNPHL 1035

RESULT 12
 PYAS MOUSE
 ID PYAS MOUSE STANDARD; PRT; 843 AA.
 AC Q91WS2; Q8KOL4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 5-like.
 GN PYPAF5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] _TaxID=10090;
 RP CONCEPTUAL TRANSLATION OF 1-287.
 RA Hinz U.;
 RL Unpublished observations (FEB-2003).
 RN [2]
 RP SEQUENCE OF 288-843 FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Rutterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
 RT "Identification of mammalian orthologs associated PYPAF5 with distinct
 functional roles.";
 RL FEBS Lett. 538:173-177 (2003).

- FUNCTION:** May mediate activation of CASP1 via ASC and promote activation of NF-kappa-B (By similarity).
- SUBUNIT:** Binds to ASC with its DAPIN domain (By similarity).
- SUBCELLULAR LOCATION:** Cytoplasmic (By similarity).

-!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).

- I - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

-!- SIMILARITY: Contains 1 DAPI domain.

-!- SIMILARITY: Contains 1 NACHT domain.

-!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.

-!- CAUTION: The N-terminus was extended using ESTs and genomic sequences in analogy to ortholog sequences

sequences, in analogy to orthology sequences.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).

EMBL; BC013519; AAH13519.1; -..

EMBL; BC031139; AAH31139.1; ALT_INIT.

MGD; MGI:2141990; Pyraf5.

InterPro; IPR001611; LRR.

InterPro; IPR007091; LRR_RNinh.

PIAM; PF00560; LRR; I.
PROCTE; PCE0934. DAPTIN. 1

PROSTIE; PS50624; DAPIN; 1.
PROSTIE: PS50837: NACHT: 1

ATP-binding: Leucine-rich repeat: Repeat.

DOMAIN	11	102
--------	----	-----

DOMAIN	168	485
--------	-----	-----

DOMAIN	569	585	ASP/GLU-RICH.
--------	-----	-----	---------------

DOMAIN	654	661	POLY-LYS.
555555	434	450	488

REPEAT	434	459
REPEAT	608	622

REPEAT	609	632
REPEAT	811	834

RELEASE	011	031
NP BIND	174	181

SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB

ry Match 11.5%; Score 868; DB 1; Length 843;

t Local Similarity 31.6%; Pred. No. 6.3e-48;

ches 249; Conservative 123; Mismatches 281; Indels 136; Gaps 22;

[illegible]

275 NQKFTQLLLQRPHPRSQDPLVKRSWPDYVEENRGHLIEIRDLF-GPGGLDTQEPRIVILQ 333

114 NKRFETKI.I.I.APGTGAVEDEI.I.GPI.GEPEPERARRSDTHTFNRI.ERGNDEESSOPI.TVVLO 173

[illegible]

334 GAAGIGKSTLARQVKEAWGRGQLYGDRFQHVYFSCRELAQSK-VVSLAELIGKDGATP 392

[illegible]

174 GPAGIGKTMAAKKILYDWAAGKLYHSQVDFAFMPCGELLERPGKRS LADLVLDQCPDRA 233

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

393 APIRQILSRPERLLFILDGVDEPGWVLQEPSEL--CLHWSQPQPADA-----LLGSLTG 445

234 WPVKRTI.AOPNRI.I.FTIDGADE---L.PTI.PSSEATPC-----KDPI.EATSGI.RVT.SGI.I.S 285

Z54 MFVNRIEAFQFNKBUFIIDGADE--UFIFDSSEAFIC-----NDFDEATSGGUKVDSGUELS 203

446 KTILPEASFLITARTALONLIPSEQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFR 505

[illegible]

286 QELPGARLLVTTTRHAATGRQLCSPQCAEIRGFSKDKKKYFFKFRDERKAERAYR 345

506 LVKSNKELWALCLVPVWSWLACTCLMQMKRKEKLTLSKTTTLCCLHYLAQALQ-AQPL 564

SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=21950691; PubMed=11821383;
RA Razmaria M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
R DiStefano P.S., Birtin J., Alnemri E.S.;
RT "CARD-8 protein, a new CARD family member that regulates caspase-1
activation and apoptosis.";
RL J. Biol. Chem. 277:13952-13958(2002).
RN [3]
SEQUENCE FROM N.A. (ISOFORM LONG).
RA Zhang H.;
RT "A novel apoptotic protein, NDP1, containing CARD and BH3 domains.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A. (ISOFORM LONG).
RA Quiet C., Vito P.;
RT "DACAR, a novel CARD-containing protein.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=21570185; PubMed=11551959;
RA Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
RA MacFarlane M., Martin S.J.;
RT "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
of multiple NF-kappa B activation pathways.";
RL J. Biol. Chem. 276:44069-44077(2001).
RN [6]
SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Kidney;
RA Guo J.H., Yu L.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
CHARACTERIZATION.
RX MEDLINE=21402909; PubMed=11408476;
RA Pathan N., Marusawa H., Krajewska M., Matsuzawa S.-I., Kim H.,
RA Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Pio F.,
RA Godzik A., Reed J.C.;
RT "TUCAN, an antiapoptotic caspase-associated recruitment domain family
protein overexpressed in cancer.";
RL J. Biol. Chem. 276:32220-32229(2001).
RN [8]
CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
RX MEDLINE=22062958; PubMed=12067710;
RA Stilo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
RT "TUCAN/CARDINAL and DRAL participate in a common pathway for
modulation of NF-kappaB activation.";
RL FEBS Lett. 521:165-169(2002).
CC -!- FUNCTION: Inhibits NF-kappaB activation. May participate in a
regulatory mechanism that coordinates cellular responses
controlled by NF-kappaB transcription factor. Involved in the
negative regulation of caspase-1.
CC -!- SUBUNIT: May form homodimers. Interacts with NEMO and DRAL. Binds
to caspase-1, pseudo-ICE and ICEBERG.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
IsoId=Q9Y2G2-1; Sequence=Displayed;
Name=Short;
IsoId=Q9Y2G2-2; Sequence=VSP_000782, VSP_000783;
CC -!- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
placenta. Lower expression in heart, kidney and liver. Also
expressed in spleen, lymph node and bone marrow.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AB023172; BAA76799.1; -.

DR EMBL; AF3222184; AAG50014.1; -.
DR EMBL; AF331519; AAK01126.1; -.
DR EMBL; AY026322; AAK08982.1; -.
DR EMBL; AF405558; AAL02427.1; -.
DR EMBL; AF511652; AAM46959.1; -.
DR InterPro; IPR001315; CARD.
DR Pfam; PF00619; CARD; 1.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS0209; CARD; 1.
KW Apoptosis; Nuclear protein; Alternative splicing.
FT DOMAIN 340 430
FT VARSPLIC 282 286
FT VARSPLIC 287 431
FT MUTAGEN 366 366
FT CONFLICT 60 60
FT CONFLICT 326 326
FT CONFLICT 422 422
SQ SEQUENCE 431 AA; 48932 MW; CB54D130807732E6 CRC64;
Query Match 8.9%; Score 668; DB 1; Length 431;
Best Local Similarity 39.5%; Pred.No. 1.7e-35;
Matches 151; Conservative 53; Mismatches 122; Indels 56; Gaps 7;
QY 1079 FWGPTGPVATEVVDKKNLYRHPFVAGSYRWPNTGLCFVMREAVTVIEFCVMDQFLG- 1137
DB 56 FLGPEGVVDVELIDKSTNRYSWFPTAGWYLSATGLGLVLRDEVTVTIAFGSWSQHLAL 115
QY 1138 EINPOHSMVAGPLLDIKAEK-GAVEAVHLPHFVALQGHVDTSLFQMAHFKEGMLLEK 1196
DB 116 DLQHEQMLVGGLPFDVTAEPPEAAVETLPHFISLQ-GEVDVSWFLVAHFQKNGMVEH 174
QY 1197 PARVELHIVLENFSPLGVLLKMHNAIFIPVTSVVLLYHRVHPPEVTFHLVLPD 1256
DB 175 PARVEPFYAVLESFSLMGLITRATSLPITSNTLIYHPHPEDIKFHLVLPD 234
QY 1257 CSIR-----KELELCYRSPGSDQ 1274
DB 235 ALLTKAIDDEDRFHGVRLQTPPMEPLNFGSSYIVSNSANKYMPKELSYRSPGIQ 294
QY 1275 LFSEPYVGHLSGIRLQVKDKDETLVWEALVKPGDMLPATLIPPARIAPVSPDAPQL 1334
DB 295 HFSFYAGQKPEIQLEITEKRHGLVMDTEVPVLDLQVAASAPP-----PFGAA- 346
QY 1335 LHFVDQYBEOLIAVTSVEVLDKLH-QQVLSQSQYERVLAEINTRPSQMRKLFSLSQSMD 1393
DB 347 --FVKENHRQLQARMGMDLQGLVDDLDQNEVLTEKEKELVEQEKTRQSKNEALLSMVEKKG 404
QY 1394 RKCKDGLYQALKETHPLIMEL 1415
DB 405 DLALDVLFRSISERDPILVSVL 426
RESULT 15
CARD_HUMAN
ID CARD_HUMAN STANDARD; PRT; 1040 AA.
AC Q9HC29; Q96RH5; Q96RH6; Q96RH8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory
bowel disease protein 1).
GN CARD15 OR NOD2 OR IBDI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND
RP VARIANT ARG-908.
RC TISSUE=Breast;
RX PubMed=11087742;

FT	VARIANT	684	684	R -> W (in CD).	QY	717	-----ESLHC-----LYETRNKTFITQV-----MAHFEEMGMCVETDMELLVC	754
FT	VARIANT	702	702	/FTid=VAR_012688.	Db	724	PAAPCEAKSVHAMPQFIWLRSLYEMQEBERLARKAAGLNVGHL-KLTFCSVGPTECAAL	782
FT	VARIANT	703	703	/FTid=VAR_012689.	QY	755	TFCIKFSRHVKLQI-----IEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSLVKVT	807
FT	VARIANT	713	713	R -> C (in CD and ulcerative colitis).	Db	783	AFVLQHLRRPVALQLDYNSVGDIGVEQLLPCLVGCKALYLRDNNISDRGICKLIECALHC	842
FT	VARIANT	725	725	R -> C (in CD).	QY	808	RNLKELDLGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTABEDCKDLAFGLRANOTL	867
FT	VARIANT	725	725	/FTid=VAR_012691.	Db	843	EOLQKALALFNNKLTDDGCAHSAWKLACRQNFU-LALRGNVITAGAQVLAEGLRGNTSL	901
FT	VARIANT	755	755	A -> V (in CD).	QY	868	TELDLSFNVLTDAGAKHLCORLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLXEL	927
FT	VARIANT	758	758	/FTid=VAR_012693.	Db	902	QFLGFMGRVGDGGAQALAEALGDHQ-SLRWLSLVGNNGSVGAQALALAKNVMLEEL	960
FT	VARIANT	778	778	A -> K (in CD).	QY	928	DIQONNLDVGVRLLCBGL-RHPACKLIRGLDQTT-LSDENRQELRALEQEKPOLLIIFS	985
FT	VARIANT	793	793	V -> M (in CD).	Db	961	CLEENHLQDEGVCSLAEGLKKNLSIKLSNNCITYLGAEAL--LQALERNDTILEVWL	1018
FT	VARIANT	843	843	/FTid=VAR_012694.	QY	986	R 986	
FT	VARIANT	853	853	E -> K (in CD).	Db	1019	R 1019	
FT	VARIANT	863	863	/FTid=VAR_012695.				
FT	VARIANT	885	885	N -> S (in CD).				
FT	VARIANT	908	908	/FTid=VAR_012698.				
FT	VARIANT	918	918	M -> V (in CD).				
FT	VARIANT	924	924	/FTid=VAR_012699.				
FT	VARIANT	955	955	A -> T (in ulcerative colitis).				
FT	VARIANT	955	955	G -> R (in CD).				
FT	VARIANT	955	955	/FTid=VAR_012701.				
FT	VARIANT	955	955	A -> D (risk factor for CD).				
FT	VARIANT	955	955	/FTid=VAR_012702.				
FT	VARIANT	955	955	G -> D (in CD).				
FT	VARIANT	955	955	/FTid=VAR_012703.				
FT	VARIANT	955	955	V -> I.				
FT	VARIANT	955	955	/FTid=VAR_012704.				
FT	MUTAGEN	305	305	K-R: NO ACTIVATION.				
FT	SEQUENCE	1040	AA; 115282	MW; 0037592D96D7DDFF CRC64;				

Search completed: January 29, 2004, 13:46:30
Job time : 26.9982 secs

QY	314	IRDLFG-PGLDQEPRIVLQGAAGIKGKSTLQARQVKEAMGRGQLYGDRFQHVYFSCREL	372
Db	278	LEELFTPGHLNDADTVLVGEGAGSKSTLLQRLHLLWAAGQDF-QEFLFVFPFSCROL	336
QY	373	-AQSKVVSIAEL-----IKDGTATPAPIRQILSRPERLLFILDGVDFPGWVLEQP	422
Db	337	QCMAPLSVRTLLFEHCWPVQVEDI-----FQLLDHPDRVLLTFDGFDE--FKFRFT	389
QY	423	SSFLCLHWSQOPADALGLSLGKTLTPASFLITARTTALQNLIPSLQARWVEVLGFS	482
Db	390	DREHCSPTDPTSQVTLNLLQGNLLKNARKVTSRPAASAFRLKYIRTEF-NLKGFS	448
QY	483	ESSRKEYFYRFTDERQAIRAFRLVKSNEKELMALCLVPVWSMLACTCLMQOMKRREKLT	542
Db	449	EQGIELYLRKRHPGVADRLIRLLOETSAHLGHLPLVFSMVSKC-----HQELLL	501
QY	543	-----TSKTTTTLCL-----HYLAQAL-----QAQPLGP-----QLRDLCSLAAGIW-----	580
Db	502	QEGGSPKTTTDMYLLILOHFLHATPPDSASQGLSPSLRGRPLTLLHLGRALWGLGMC	561
QY	581	-----OKKTLFSPDDLKRGHGLDCAIISTFLKMGILQEHPIPLSYSTHLCFOEFFAA	632
Db	562	CYVFSAQQLQAAQVSPDDISL-----GFLVRA-----KGVVPGSTAPL--EFLHITFOCFFAA	612
QY	633	MSYVLED-----EKGRGKHSNCIIDLKTLLEAYGI	662
Db	613	FYLAISADVPALLRHLNCRPGNCPMARLLPTWCIOASEGKSSVAALLQKA-EPHNL	671
QY	663	HGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNL-----MQVPSLQLLLQPHSL-	716
Db	672	Q-----ITAAFLAGLLSRE---HWGLLABECQTSKALLRRQACARWCLARSRLKHFHSIP	723

Query Match 5.2%; Score 393.5; DB 1; Length 1040;
Best Local Similarity 25.4%; Pred.No. 2.8e-17; Indels 147; Gaps 31;
Matches 198; Conservative 111; Mismatches 325; Indels 147; Gaps 31;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 79.1934 Seconds
(without alignments)
4656.416 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534

Sequence: 1 MAGGAWRLACYLEFLKKEE.....HLIMLWKGSKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_protein.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rviro.*

16: sp_bacterioph.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2640	35.0	509	4	Q9H5Z8	Q9h5z8 homo sapien
2	1754	23.3	409	4	Q9H5Z7	Q9h5z7 homo sapien
3	1371.5	18.2	892	4	Q8NEU4	Q8neu4 homo sapien
4	1254.5	16.7	1033	11	Q8R4B8	Q8r4b8 mus musculus
5	899.5	11.9	982	11	Q8BU40	Q8bu40 mus musculus
6	868	11.5	863	11	Q8CEJ9	Q8cej9 mus musculus
7	843	11.2	846	4	Q8IXT0	Q8ixt0 homo sapien
8	815.5	10.8	713	6	Q95LZ7	Q95lz7 macaca fasc
9	689.5	9.2	748	11	Q99MW0	Q99mw0 mus musculus
10	679.5	9.0	825	11	Q8C6M5	Q8c6m5 mus musculus
11	676	9.0	657	11	Q8EPG7	Q8epg7 mus musculus
12	573	7.6	673	11	Q8CCN1	Q8ccn1 mus musculus
13	393.5	5.2	953	4	Q8IWF5	Q8iwf5 homo sapien
14	392	5.2	1020	11	Q8K3Z0	Q8k3z0 mus musculus
15	383	5.1	287	4	Q9BY26	Q9by26 homo sapien
16	372.5	4.9	778	4	Q8NPF48	Q8nfp48 homo sapien

ALIGNMENTS

RESULT 1

Q9H5Z8 PRELIMINARY; PRT; 509 AA.

AC Q9H5Z8; DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 23, Last annotation update)

DE Hypothetical protein FLJ2740.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]_TaxID=9606;

RP SEQUENCE FROM N.A.

RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,

RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,

RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,

RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,

RA Isogai T., Sugano S.;

RT "NEDO human cDNA sequencing project.;"

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK026393; BAB15469.1; --

DR HSSP; PI0775; 2BNH.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007091; LRR_RNinh.

DR Pfam; PF00560; LRR; 1.

DR PROSITE; PS50503; LRR_R1; 2.

KW Hypothetical protein.

SQ SEQUENCE 509 AA; 56908 MW; EB7535AF69817F5B CRC64;

Query Match 35.0%; Score 2640; DB 4; Length 509;

Best Local Similarity 100.0%; Pred. No. 2.2e-190;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 735 MAHFEEMGCVETDMLLVCTCIKFSRHVKQLIEGRHSTWSPVWVLFWRVPTD 794

Db 1 MAHFEEMGCVETDMLLVCTCIKFSRHVKQLIEGRHSTWSPVWVLFWRVPTD 60

Qy 795 AYQWILFVLKVRNLKELDLGSLSHSAVKSCLKTLRRPRCLLETTLRAGCGLTAEDC 854

```

Db 61 AYQILFVSVLKVTRNLKELDLGNSLSHSAVSKLCTLRPRCLLETLLRAGCGLTAEDC 120
QY 855 KDLAFGLRANQTLTDLDFSNVLTDAAGKHLCORLPORPCKLORLQVSCGLTSDCCODL 914
Db 121 KDLAFGLRANQTLTDLDFSNVLTDAAGKHLCORLPORPCKLORLQVSCGLTSDCCODL 180
QY 915 ASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRLGLDQTLTSDMRQELRAL 974
Db 181 ASVLSASPSLKELDLQNNLDDVGVRLLCGLRHHPACKLIRLGLDQTLTSDMRQELRAL 240
QY 975 EQEKPQLLIFSRKPKSVMTPTGLDGTGEMSNSTSLKQRLGSGRAASHVAQANLKLDV 1034
Db 241 EQEKPQLLIFSRKPKSVMTPTGLDGTGEMSNSTSLKQRLGSGRAASHVAQANLKLDV 300
QY 1035 SKIFPIAETAESSPEVPVELLCVPSASQGLHTKPLGTDDDFMGPTGPVATEVDKX 1094
Db 301 SKIFPIAETAESSPEVPVELLCVPSASQGLHTKPLGTDDDFMGPTGPVATEVDKX 360
QY 1095 KNLRYRHFPVAGSYRPNGLCFVMEAVTVEIFCVDQDFLGEINPQHSMMVAGPLLDI 1154
Db 361 KNLRYRHFPVAGSYRPNGLCFVMEAVTVEIFCVDQDFLGEINPQHSMMVAGPLLDI 420
QY 1155 KAEPGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLEKPARVELHHIVLENSPSP 1214
Db 421 KAEPGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLEKPARVELHHIVLENSPSP 480
QY 1215 LGVLLKMIHNALRFIPVTSVLLYHR 1240
Db 481 LGVLLKMIHNALRFIPVTSVLLYHR 506

RESULT 2
Q9H527 ID Q9H527 PRELIMINARY; PRT; 409 AA.
AC Q9H527;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22745.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Obayashi M., Nishi T., Shibahata T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026398; BAB15470.1; -.
KW Hypothetical protein.
SQ SEQUENCE 409 AA; 46068 MW; 07CC5FACF3EB7236 CRC64;

Query Match 23.3%; Score 1754; DB 4; Length 409;
Best Local Similarity 87.7%; Pred. No. 1.le-123;
Matches 343; Conservative 0; Mismatches 0; Indels 48; Gaps 2;

QY 967 MRQELRALQEKPOLLIIFSRKPKSVMTPTGLDGTGEMSNSTSLKQRLGSGRAASHVAQ 1026
Db 1 MQELRALQEKPOLLIIFSRKPKSVMTPTGLDGTGEMSNSTSLKQRLGSGRAASHVAQ 60
QY 1027 ANLKLDVSKIFPIAETIA-----EESPVEVPVELLCVPSASQGLHTKPLGTDDDFWGP 1082
Db 61 ANLKLDVSKIFPIAETIAKSHSESSPEVPVELLCVPSASQGLHTKPLGTDDDFWGP 120
QY 1083 TGPVATEVDKKNLRYRHFPVAGSYRPNGLCFVMEAVTVEIFCVDQDFLGEINPQ 1142
Db 121 TGPVATEVDKKNLRYRHFPVAGSYRPNGLCFVMEAVTVEIFCVDQDFLGEINPQ 180

```

```

QY 1143 HSMVAGFLDDIKAPGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLEKPARVEL 1202
Db 181 HSMVAGFLDDIKAPGAVEAVLPHFVALQGGHVDTSLFQMAHFKEEGMLEKPARVEL 240
QY 1203 HHIVLENSPSPGLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFFHLYLIPSDCSIR-- 1260
Db 241 HHIVLENSPSPGLGVLLKMIHNALRFIPVTSVLLYHRVHPEEVTFFHLYLIPSDCSIRKA 300
QY 1261 -----KELELCVRSPEGDQLFSE 1278
Db 301 IDDLMEKQFVRIHKPPPLTPLYMGCRTVSGSGMLEILPKLELELCVRSPEGDQLFSE 360
QY 1279 FVVGHLGSGIRLQVKDKKDETLVWEALVKPG 1309
Db 361 FVVGHLGSGIRLQVKDKKDETLVWEALVKPG 391

RESULT 3
Q8NEU4 ID Q8NEU4 PRELIMINARY; PRT; 892 AA.
AC Q8NEU4;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Monarch-1 splice form IV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams K.L., Linhoff M.W., Ting J.P.Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY116207; AAM75145.1; -.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTFase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50503; LRR_R1; 1.
DR PROSITE; PS50837; NACHT; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

Query Match 18.2%; Score 1371.5; DB 4; Length 892;
Best Local Similarity 35.2%; Pred. No. 2.6e-94;
Matches 356; Conservative 135; Mismatches 349; Indels 171; Gaps 22;

QY 8 RIACYLEFLKKEELKEFOLLANKAHSRSSGETP-AQPEKTSGMVASYLVQYGEORA 66
Db 12 RLSTVLELEAVELVKKFKYLGTA--TELGEKIPWGSMEKAGPLEMAQLLITHEGPEEA 69
QY 67 WDLALHTWEQMLRSLCAQAQEGAGHSFPYSSEPHLGSFSQ-PTSTAVLMPHWHLP 125
Db 70 WRLALSTFERINRKLWERGQREDLVRDTPPGGSS--LGNQSTCLLEVSIVLTP--RKDP 125
QY 126 AGCTQGSERRVRLQPLDPTSGRRWRREISASLLVQALPSSPDHESPSQESPNAPTSTAVLGS 185
Db 126 QETRDYRVRFRLEMDRNARLGCVCNLSHRYTRLLLVKEHSNPMQVQ----- 173
QY 186 WGSPPQPSLAPREQAPGTOWPLDTSIGIYYTEIREREREKSEKGRPPMAAVVGTPPOAH 245
Db 174 ----- 173
QY 246 TSLQPHHPWPSPVRESLCTWPKWKNEDFNQKFTQLLLQLRPHRPSQDPLVKRSHWPDYVE 305
Db 174 -----QQLLDTGRGHARTV----- 187
QY 306 ENRGHL---IBIRLFGGLDQOE--PRIVILQGAAGICKSTLARQVKEAWGRGLYGDGF 361
Db 188 ---GHQASPKIETLFEDEBERPPRTVVMQAGIAGICKSLAHKVMLDWADGKLFGQGRF 244
QY 362 QHVFFYSCRELAQSKV-VSLAELIKGQGTATPAPIQTLSRPERLLFLDGVDEFGWVLIQ 420

```

Db 245 DYLFYINCREMNASQTECQMDLIFSCWPEBPAPQLRIVPERLLFIIDGDELDKPSPH 304
QY 421 EPSSELCHWSQPQADALLGSLGKTLILPEASFLITARTTALQNLIPLEQARWVEVLG-480
Db 305 DQGPWCLWEERKPTPELLNSLRKLPPELSTLITRPTALEKHLRLEHPRHVEILG 364
QY 481 FRESSRKEYFYFTDERQAIARFRLVKNKELMALCLVPWVSWLACTCLMOQMKRKEKL 540
Db 365 FSEARKEYFYFHNABAGOVFNVRDNEPLFTMCFVPLVWVWVCTCLOQLEGGLL 424
QY 541 TLTSTKTTTTLCHLYLAQALQAPLGPOL-----RDLSLAAGTGWOKTLFSPDRLKH 594
Db 425 RQTSRTTAVMYDYLUSLQPKPAPRLOPPNQRGLCSLADGLMNQKILFEEQDLRKH 484
QY 595 GLDGAIISTFLKMGILQEH-PIPLSYFHLFCQEFFAAMSYVL-EDEKRGKHSNCIID 652
Db 485 GLDGEDVSFAFNWIFQKDINCERYYSFHLSPQEFFAAMSYVILDBEGEGAPDQ----D 540
QY 653 LKLTLEYGI-HGLFGASTTRFLGLLSDEGEREMENIFCHRLSQ--GRNLQMWVPSLO- 708
Db 541 VTRLLEYAFSRSFALTSRFLGLNBEETRSHLEKSLCWKVSPIHKMDLLQWISKAQ 600
QY 709 ---LLIOPHSLESLHCLYETRNKTLTQVMAHPEEMGMC-VETDMELLYCTCIKFSRHV 764
Db 601 SDGSTLQOGLSEFFSCLYEIQEEFTQALSHFPQIVIVSNIAAKMEHMYVSFCLKRCRSA 660
QY 765 KKLQLT-----EGRQHRSTWSPWVLF--WVPVTDAYWQIILFSLVKVTRNLKELD 814
Db 661 QVHLHYGATYSADGEDARCSAGAHTLLVQLRPVTLVDAYSEHLAAALCTNPNIELS 720
QY 815 LSGNSLSHSAVSKLCTLRPRCLLETLRLAGCGLTAEDCKDLAFLGRANQTLTDLDSF 874
Db 721 LYRNALSGVGLKLCQGLHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 780
QY 875 NVLTDAGAKHLCLORLPQSPCKLQRLQVSCGLTSDCCOLASVLSASPSIKELDLQONML 934
Db 781 NGVGFPMMLLCEGLRHPQCRQMIQLRKQCESGACQEMASVLTGNPHLVLDLTGNAL 840
QY 935 DVGVFLLCEGLRHPACKLRL-----GLDQTLSDMRQELRALEQEKPOL 981
Db 841 EDGLRLKLCQGLRHPVCRLLTWLFGMD---LNMKTHSRLAALRYTKPYL 888

RESULT 4
ID Q8R4B8 PRELIMINARY; PRT; 1033 AA.
AC Q8R4B8;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Mast cell maturation inducible protein 1.
GN MWIG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.,
RT Identification of inducible genes during in vitro maturation of mouse
RT bone marrow-derived mast cells to connective tissue-type mast cells."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486632; AAL90874.1; -
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001611; LRR
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.

DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_RI; 1.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS0503; LRR_RI; 3.
DR PROSITE; PS05037; NACT; 1.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B12117 CRC64;
Query Match 16.7%; Score 1254.5; DB 11; Length 1033;
Best Local Similarity 32.6%; Pred. No. 2.e-85;
Matches 331; Conservative 164; Mismatches 334; Indels 187; Gaps 23;
QY 8 RIACVLEFLKKEELKEFOLLANKAHSRSSSETPAQPEKSGMEVASVLYAQYGEORAW 67
Db 7 KLAQYLEDLEDVLLKKFKHLEDYPPKGCIPVPRGQMEKADHLDLATIMDFNGEEKAW 66
QY 68 DLALHTWQMGLRSICAAQAGAGHSPSPFPSPHLSGSPQSTSTAVLMPWIELPAG 127
Db 67 ANAVMIFAINRRDLWEKAK-----DQP-----EWNDT 95
QY 128 CTQGSERRVLRQLPTSGRRWRREISASLLYQALPSPDHPSPQSPNAPTSTAVLGSWG 187
Db 96 CTSHSS-WVQEE--DSLSEEW-----MGLLG 118
QY 188 SPQPSLAPREQEPGTQWPLDETSGIYYTEIRERERKSEKGRPPMAAVGTTPQAHTS 247
Db 119 YLSRISICKKK-----DYCKMYRRHVRSRFYSIKDRN-----ARLG----- 155
QY 248 LOPHHHPPEPSVRESLCTWPKNEDFNQKFTQLLLQRPHPRSQDPLVKRWPDPYVEN 307
Db 156 -----ESVDLSRYTQLQLV-KEHPSKQE----- 178
QY 308 RGHLL-----EIRD-----LFGPGLDTQP-RIVILQGAAGIGKSTLAROVKEA 350
Db 179 REHELLTTRTKMRDSPMSSLKLELLFPEDGHSPEVHTVFGAAGIGKTLARKIMLD 238
QY 351 WGRGOLYGRFOHVFYFSCRELAQSKVSVLAEIKGDTATPAPRIQLISRPERLLFILD 410
Db 239 WALGKLFKDFLYLFFHCREVSLRTPSLADLVSCWDPDPNPVCKILRKPSRILFLMD 298
QY 411 GYDEFGWVLOEPSSELCHWSQOPADALLGSLGKTLILPEASFLITARTTALQNLIPSL 470
Db 299 GFDELQAGFDEHIGEVCTDWQKAVRGDILLSSLRKLLPKASLLITTPVALEKLOHLL 358
QY 471 EQARWVEVLGSPESRKEYFYFTDERQAIARFRLVKNKELMALCLVPWVSWLACTCL 530
Db 359 DHPRHVEILGFSEARKEYFYFYNELQARAFRLIQENEVLFTMCFPLVCIWVCTGL 418
QY 531 MQQMKRKEKLTLSTKTTTTLCHLYLAQALQAP-----LGPOLRDLCSLAAGIWKQKT 584
Db 419 KQOMETGKSLAQTSTKTTTAVVVFLLSLQSRGGIEEHLFSDYLOGLCSLAADGIWNQKI 478
QY 585 LFPDOLRKHGLDGAIIITFLKMGILQ-EHPTPLSYFHLFCQEFFAAMSVLEDE---- 640
Db 479 LFEEDCLRHGLQKTDVSAFLRMNVFQKEDCERFYSFSHMTFQBFFAAMYLLLEEARG 538
QY 641 -----KGRGKHSNCI-IDLEKTEAYG-IHGLFGASTTRFLGLLSDEGEREMENIFCHRL 694
Db 539 ETVRKGPGGCDLLNRDVKVLLNENTKFKEGVLIIVFVRFGLVNGQRTSYLEKKLSCKI 598
QY 695 SQ--GRNLQMWV-----PSLQLLQPLSHLSLHCLYETRNKTLTQVMAHPEEMGMCVETD 748
Db 599 SQQVRLLELLKWLVEKAKAKKLOQWQSLELFCLYEMQEEDFVQSNMDHFKPIELNLSR 658
QY 749 MELLVCTFCIKFSRHVKKLQ-----IEGRQHRSTWSP-----MVVLFRW 790
Db 659 MDHVVSFFCIKCHRVKTLISGLGFFHNSPKEEBERRGGRPLDQVQCVFPDTHVACSSRLV 718
QY 791 P--VTDAYWQIILFSLVKVTRNLKELDLSONSLSHSAVKSLCKTLRPRCLLETTLRAGCG 848
Db 719 NCLTSSPCRGULFSSLTNRSRLTDLSDNTLGDPMGRVLCALQHPGNCIQRMLGRGC 778
QY 849 LTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLCLORLPQSPCKLQRLQVSCGLTS 908

```
Db 779 LSHQCFFDSSVLSQKVELDLSNALGDFGIRLLCVGLKHLNQLKWLVSCLTS 838
Qy 909 DCCODLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHPACKLRRLGLDQTLIS 964
Db 839 ACCQDIALVLSNHNSTRLYIGENALGDSGVQLCEKMKDPQCNQKLGVLNSGLT 894

RESULT 5
Q8BU40
ID Q8BU40 PRELIMINARY; PRT; 982 AA.
AC Q8BU40;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Similar to PAN2 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK087843; BAC40024.1; -.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFEA2053585 CRC64;

Query Match 11.9%; Score 899.5; DB 11; Length 982;
Best Local Similarity 26.8%; Pred. No. 1.2e-58;
Matches 269; Conservative 176; Mismatches 366; Indels 191; Gaps 26;

Qy 203 GTQWPLDSTGIYTBIRE-REREKSEKG-RPPWAAVVGTPPQAHTSLQPHH-----PW 255
Db 9 GLMWYLEELNKEFVKFKEFLQVQLGKQVSWTEVKKASQDLASLLXHYEKPAP 68
Qy 256 EPSVRESLSTWPKNEDNQKPTQLLLORPH-----PRSQDPLVKRSDPYVEENRG 309
Db 69 DMTFR-----PFQINRMDLIERAKREIDGCPKLYRAHMKTKMTH--DSSRA 113
Qy 310 HLIEIR-----DLFGPLDQEPRIVLQGAAGICKSLARQVKEAMGR 353
Db 114 FTISQNFLEKPTEDDYCFENLFOSKTESKPQVFLSGGAGVGTLMKRLMLAWTE 173
Qy 354 GOLYGRFOHFVYFSCRELAQKVVSLAELIGKDGATAPAPIRQILSRPERLLFILDGVD 413
Db 174 SPVFLHKFSYIFVCCREVKQLKTASLAELISREWPGSPAPIEELISKPEKLLFIIDSL 233
Qy 414 EPGWLVQEPSELCLHWSQFQADALLGSLGKLTILPEASFLITARTTALQNLIPSLEQA 473
Db 234 GMECDLFKWESELCDNCTEQPNVLLSLLARLKPESLLISATPSEFKEMENRIEYT 293
Qy 474 RWVEVLGFSERSEKVFYFYFDTQRAIRAFRLVKSNEKLMALCLVPWVSWLACTCLMQQ 533
Db 294 HVKIKGLKERNIKOMSFHRLFQDRNRAHEAFSLVRENEQLFTVCQVPVLCWVATCLKEE 353
Qy 534 M-KRKEKLTSTKTTTTLCHVLAQALQALP-----QLRDLCSLAAGIWKQKTLF 586
Db 354 IEKGRDPVICRCTSLYTHFNLFIPQNAHSPKKSODQLQGLCSLAEGWTTDTFFV 413
Qy 587 SPDDLAKHGLDGAISTFLKMGIL-QEHPPIPSYFIHLCOFEFFAAMSYYVLEDEKGRGK 645
Db 414 GEEALRRNGIMSDIPILLDIGMLINIRESEKSYIFLHPSVQEVCAAI FYLL--KSHVD 470
Qy 646 H-SNCIIDLKTELEAY-----GTHGLFGASTTFLGLLSDEGEREMENIFHCRLSQ--GR 698
Db 471 HPSQEVKSIKLMFAFLKVKVQWIFFGS---FIFGLHESQKLEAFHGLSQEIKR 527
Qy 699 NLQWVPSL-----QLLLQPHSLHSLHCLYETRNKLTPLQVMAHFEBMGMCVETDMELLVC 754
```

```
Db 528 QLYQCLETISGNEELQEQIDGMKLFYCLFEMDDDTFLVEAMNMEQINFVAKDYSDIVA 587
Qy 755 TFCIKFSRHVKL-----QLIEGRQHS-----TWSPTWVL-----FRWVPV 792
Db 588 AHCLKHCFTLKLSFSTQNVLSGAQHSYMERLLTAMNHICSVFIISKDIQELRMKDTNL 647
Qy 793 TDAYWQILFSLVKVTR-----NLKELDLSGNSLSHSAV 825
Db 648 SESAFSVLYNNLYHNYTLNVIVANNVFFVCEKYLFFELIQNCNLQHLNLSLTILSHSDV 707
Qy 826 KSLCKTLRRPRCLLETTLRACGLTAEDCKDLAFGLRANQTLTDLDFNVLTDAKHL 885
Db 708 KLLCDVLSQAECNIBELVVAACSISSDDCKVFASVLISNKTLLKHLNLSNL-DKGIASL 766
Qy 886 CORLRQPSCKLQRLQVSCGLTSDC-----NLDKELDLSGNSLSHSAV 910
Db 767 CKSLCHPCILHLVLANCSLNEKCDVLSVLRNRNKTLSHLDISSNDLKDEGLKVLQRA 826
Qy 911 -----CQDLASVLSASPSLKELDLQNNLDDVGVRLLCGLR 948
Db 827 LTPDSVLKSLRHLCLITISCCQDLAEVLRNNQNLVLSQVSNKLEDTGVKLLCDAIKH 886
Qy 949 PACKLIRLGLDQTLTSDMRQELRALEQKPOLIFSRKPSVMTPTGLDTGEMSNSTS 1008
Db 887 PNCHELDLGLACELTGACCEDLASTFTQCKTLWAMNLLK-----NALDYNGLVLCE 939
Qy 1009 SLKQRLQSERAAASHVAQANIKLLDV-SKIPPIAIEAESPP 1049
Db 940 ALKQO-----TCATYVLGLQITDFDTETQAFVLAEE--QEKNP 974

RESULT 6
Q8C6J9
ID Q8C6J9 PRELIMINARY; PRT; 863 AA.
AC Q8C6J9;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Hypothetical RNI-like structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK054426; BAC35775.1; -.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 99955 MW; D3A46C782DA544CF CRC64;

Query Match 11.5%; Score 868; DB 11; Length 863;
Best Local Similarity 27.7%; Pred. No. 2.4e-56;
Matches 236; Conservative 154; Mismatches 343; Indels 120; Gaps 17;

Qy 214 IYVTEIRERE-----KSEKGRPPWAA-----VVG---TPQAHTSLQPHHPWP 257
Db 41 ISWTEVREASREDLAILLVKHCQGNQAWDTTFRVQMIGRNVITNRAETGAHSTIYRA 100
Qy 258 SVRESLSTWPKNEDFNQKFTQLLLQRPSPQDPLVKRSDPYVEENRGHLIEIDL 317
Db 101 HLKEKL-----THDCSRKE-----NISQNFQDEVDHL---ENL 132
Qy 318 FPGPLDQEPRIVLQGAAGICKSLARQVKEAMGRGOLYGRFOHFVYFSCRELAQSKV 377
Db 133 LVPNGTENNPNKVVVLQGVAGIGKTIKLLKMLIVMSEGLVFNQKFSYIFVFCCHQVQLQ 192
```



```
QY 378 VSLAEILGKDGATAPAPIRQILSRPRLFLILGVDVPCGWVQEPSSSELCHWSQPQAD 437
Db 193 ASLADLISREWSPSPAMEILLSQEKLFIIDSLEGMWVNTQQDSQLCYNCMEKQPVN 252
QY 438 ALLGSLGKTLIPASFLITARTALONLIPSLQARWVEVLGFSSESSKEYFYRYFTDE 497
Db 253 VLLSSULRKLIPESSELLSTSCETPKDKMDWIEYTNVNTTIGFKENNINMCFHSLFQDR 312
QY 498 ROAIRAFRLVKSNEKELWALCLVPWSWLACTCLMQOMKR-KBKLTULTSTTTTCLHLYLA 556
Db 313 NIAQEAFLIRENEQLFTVCQAPVVCVMVATCLKNEIESKOPVSICRETTSTLYTHILN 372
QY 557 QALQAOQLPGQ-----LRDLCSLAAGIWOKTLFSPDDLRKHGLDGAIISTFLKMGIL 610
Db 373 LFIPHNAQPNNSNSEDLLNLCFLAVEGMWTDISVNEEALRRNGIMDSIPTLLDIGIL 432
QY 611 -QEHPIPLVSFTHLCFOFFAAMSVLDEKGRGHSNCID-----LEKLEAYGIH 663
Db 433 EQRESSENSYIFHPVSQVQFCAMFYLHSEMDHSCQGVYFETFLFTFLNKIKKQWVFL 492
QY 664 GLFGASTTRFLGLLSDGEGEREMENIFHCRLSQ--GRNLMQWVPSLQLLQP-----HSL 716
Db 493 GC-----FFGGLLHETGEKLEAFGGYHLSKELRRQLFLMLELLDTHLPDVKKINTM 545
QY 717 ESHLCYETRNKTLFQVMAHPEMGCMVETDMELLVCTFCIKF-----SRHVKK 766
Db 546 KFFYCLFEMEEVFGQSANCREQIDVWVGYSDFVAAVCLSHGSAITDFSTSAQNVLN 605
QY 767 LQIEGRQHSRSTWPTMVFLRVVPV-----TDAYWQLFLSVLK-----805
Db 606 BELGQRKULLIHLQICSVFLRNKDKTLRIEDTIFNEPVFKIFYSLYKNSCILKTLVA 665
QY 806 -----VTRNLKELDLSGNSLSAVKSLCKTLRRPRCLLETLRLAGCGL 849
Db 666 YNVSFCLDKKFLFELIQSYNLELYLGRFTLSHSDVEMLCIDILNQECHIRILDANCSL 725
QY 850 TABDKDLAFGLRANQTLTDLDSFNVLTDAQAKHLQORLQPSCKLQRLQVSCGLTSD 909
Db 726 CEHSWYLSVLRQNSLRVYNISYNNLKDEGLKALCRALTLPNSALHSLSLEACQLTGA 785
QY 910 CCODIASVLSASPSLKELDLOQNNLDVGVRLICEGLRHCPACKLIRGLDOTTLSDEMQR 969
Db 786 CKDOLASTFTRYKCLRRINAKNSGSLGFLVLCRAKQDQOTCTLYELKLRMADF-DSDSQ 844
QY 970 ELRALQEQPOL 982
Db 845 EPLLSEMERKIL 857
RESULT 7
Q8IXT0 PRELIMINARY; PRT; 846 AA.
AC Q8IXT0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE Similar to NALP2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039269; AAH39269.1; -
SQ SEQUENCE 846 AA; 96368 MW; 46BE8245550E39F7 CRC64;
Query Match 11.2%; Score 843; DB 4; Length 846;
Best Local Similarity 30.0%; Pred. No. 1.8e-54;
Matches 255; Conservative 154; Mismatches 328; Indels 114; Gaps 28;
```

RESULT 8

```
Q95LZ7
ID Q95LZ7 PRELIMINARY; PRT; 713 AA.
AC Q95LZ7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 80.3 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
```

```

OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB071044; BAB64437.1; -.
DR InterPro: IPR007091; LRR RNinh.
DR PROSITE: PS50503; LRR RI; 2.
DR PROSITE: PS50837; NACHT; 1.
KW Hypothetical protein.
SQ SEQUENCE 713 AA; 80258 MW; 6F214C9B773F54DC CRC64;

Query Match 10.8%; Score 815.5; DB 6; Length 713;
Best Local Similarity 31.4%; Pred. No. 1.6e-52;
Matches 228; Conservative 114; Mismatches 248; Indels 135; Gaps 18;

QY 383 LICKDGTATPAPRIQLSRPERLLFILDGVDE----PGWVLQEPSELCLHWSQOPADA 438
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MVRFDWPELQDDIPHLAQAKILFVIDGFELGAPPGALIQ----DICGDWEQKPEVPV 56
QY 439 LIGSLLGKTLTPASFLITARTALONLIPSLQARWVEVLGFSSESRKEYFYRYFTDER 498
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 LIGSLLRKRLPKATLVVTRPALRDRLFLABQPIYIRVEGFLEDRRAYFLRHFGDED 116
QY 499 QAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKLTLSKTTTTLCHLVL-AQ 557
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 QAMRAFELMRSAALFQLGAPAVCVICTTLKLQMEKGEDPAPTCLTSTGLFLRLCSQ 176
QY 558 ALQAPLGLQRLDCLSLAAGIWQKTLFSPDDLRKGLDGLAIISTIFLKGILQHPPLP 617
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 FPQGAQLRGALRALSLAAQSLWAQMSVLHGDELSAGVQESDLRLFLDGLDILQDGVAK 236
QY 618 S-YSFTHLCFOEFPAAMSIVL--EDSKGKGKHSNCIIIDLEKTL--EAYGIHGLFGASTT 671
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 GCYSFTHLSFQQLFTALFALEKEEDEDGDHAWDIDGVQKLSREERLKNPDLQAG-- 294
QY 672 RFLGLLSDEGEREMENI FHCRLSQGNLWQWVPSL-QLLQ-----PHSLES 718
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 RFLFLGANERVKLEATFGWRMS-----PEIKQELLRCDSVRKNGHFTAADLREL 345
QY 719 LHCLYTRNKTFLTQMAHPEEMGCVETDMELLVCTFCIKESRHV-----KKLOJ--- 769
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 LCCLYESQDEDLVKEYMAQFKEISHLNA-VDIAPSSFCFKHCQNIFPADACRNLCALR 404
QY 770 -----IEGRHRSWSPTMVVL-----PRWV-----PVTDAYWQLLFSVLKVTEN 809
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 GHKVTHTLIQGDQDMDLPALCEVLRHPECNRYLGLVSCSNATTQWADLSALEANRS 464
QY 810 LKELDLSGNSLSHSAVSKLCTLRPRCLLETLRAGCGLITABDCKDLAFGLRANOTLTE 869
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
465 LMCVNLSDNELDEGAKLLYTLRHPKFLQRLSLENCHLTRANCKDLAAVLVWSRELTH 524
QY 870 LDLSFNVLTDAGAKHLCORLQPSCKLQRLQVSCGLTSDCCODLASVLSAPSCLKLDL 929
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 LCLAKNSLKDVTGVKFLCEGLSYPECKLQALVWNCDDITSDDGCSLAKLQESKSLCLDL 584
QY 930 QQNNLDVGRVLLCEGLRHPACKL-----953
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
585 GLNHIGVTGVKLCEALSPKLNRLCWLWGCISIPFSCEDLCSALSCNQSLSTLDLQGN 644
QY 954 -----IRLGLDQTTLSDEMRELRALEQKPOLLIFFS-----985
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
645 PLGSSGVKMLFKLTTRPGTLQTLRLKIDD--FNDELHKLLEETEENNPQLIIDTEKHPDW 702
QY 986 RRPKS 990
DB : : : :
703 KRPKS 707

```

```

RESULT 9
Q99MW0 PRELIMINARY; PRT; 748 AA.
AC Q99MW0;
AT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribonuclease/angiotensin inhibitor 2.
GN RNH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525;
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogonia.";
RL Nat. Genet. 27:422-426(2001).
DR EMBL: AF285581; AAK31960.1; -.
DR HSP; F10775; 2BH.
DR MGD; MGI:1890518; Rnh2.
DR InterPro: IPR007091; LRR RNinh.
DR PROSITE: PS50503; LRR RI; 1.
SQ SEQUENCE 748 AA; 84946 MW; 674A385E011DE8EC CRC64;

Query Match 9.2%; Score 689.5; DB 11; Length 748;
Best Local Similarity 26.2%; Pred. No. 5.6e-43;
Matches 197; Conservative 120; Mismatches 241; Indels 193; Gaps 14;

QY 419 LOEPSSELCHWSQOPADALLGSLGKTLTPASFLITARTALONLIPSLQARWVEV 478
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 MSERESELCDTCTEQPURLISLLRRMLPKSFLISATPEFKMEGRVECTNNKIV 64
QY 479 LGFSSSRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQOM--RRK 537
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 TGFNESNKMVFRSLFQDKTKTQEIFSLVKENQQLFTVCQVPLCMVATCLKKEIKGR 124
QY 538 EKLUTSTKTTTLCUH-----YLAQALQ--AQLGPDRLDCLSLAEGIWQKTLFSPDDL 591
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 DLVSVCRRRTSLYTTTHIFNFIPOSAQYPSKESQAQLQSLCLSLAAEGMWTDTFVFEAL 184
QY 592 RKHGDLGAIISTFLKMGILQ--HPIPLSYSTHLCFOEFFAAMSIVLEDEKGRKHSNCI 650
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 RRMGIMDSIDPILDLVRILEKSKSEKSYIFLHPSIQEVCAIFYLL---KSHMDHPQD 241
QY 651 IDLEKTLA-----YGIHGLFGASTTRFLLGLSDSEGEREMENIFHCRLSQ--GRNL 700
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
242 V---KSIEALFTFLKKVQWIFPGS---FIFGLHSEQKLEAFFGHQLSQBIKRLQ 295
QY 701 MQWVPSL-----QLLQPHSLSLHCLYETRNKTLFTQMAHPEEMGCMVETDMELLVCTF 756
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 YQCLTISGNEELQVQDMKLFYCLFEMDDDEAFQAAMNCMEQINFVAKDYSDVIVA 355
QY 757 CIKFSRHVKLQ-----LIEGRHRSWSPTMVVL-----953
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 CLQHCSTLKLKLSLSTQNVLSQEQHSYTEKLIMCWHHMCVLSISKDIYIIVQKNTLN 415
QY 780 -----979
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 TASLVLYSHLMYPSCTLAKALVNNVTVFLCDNRLFPFELIQNCLQHLDLNLTLFSLHGDVKL 475
QY 780 -----SPTMVVLFRWV-----790
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
476 LCDVLSQBECEKIEKLWAAACNLSPDDCKVFASVLISSKMLKHLNLSNNLDKGTSLSKA 535
QY 791 -----PVTDAYWQLLFSVLKVTENKELDLSGNSLSHSAVSKLCTLR 834
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
536 LCHPCVILKNLVLNCSUSEQCWDYLSSEVLRNRKNTLNHLDISSNDLKDEGLKVLCLALSL 595
QY 835 PRCLLETLRAGCGLITABDCKDLAFGLRANOTLTLELDLSFNVLTDAGAKHLCORLQPSCK 894

```

```
Db 596 PSVLKISVRYCLITTSQCPLAEVLRKNQNLRLQVSNKIEDAGVKLLCDAIKHNC 655
Qy 895 KIQRLQVSCGTLTSCODLASVLSASPSLKELDLQNNLDVGVRLICEGLRHHPACKLI 954
Db 656 HLENGICLACALTGACCEDLASAFTHCKTLWGINLQENALDHSGLIVLFEALQKQOCTLH 715
Qy 955 RIGLDQTLSDMRQELRALEQKPOLIFS 985
Db 716 VLGLRITDFDKETQELLMAEEKNPHLSILS 746

RESULT 10
Q8CGM5 PRELIMINARY; PRT; 825 AA.
AC Q8CGM5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Weakly similar to PAN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL: AK054264; BAC35710.1; -.
SQ SEQUENCE 825 AA; 94021 MW; 11D71DEABEAFCA31 CRC64;

Query Match 9.0%; Score 679.5; DB 11; Length 825;
Best Local Similarity 26.0%; Pred. No. 3.7e-42;
Matches 212; Conservative 150; Mismatches 313; Indels 141; Gaps 24;

Qy 208 LBETSIIYTEIREREREKSEKRPWAAVGTTPQAHTSLOPHHHHPWPSRESLCSW 267
Db 76 LBEINRVDLAELVVKQIEE-----AVLVKPEKVS-----KPREPSG-----TLTF 117
Qy 268 PKNEDFNQKFTQLLLQRPSPDPLVKR-----SWPD-----YVENRGHLIE 313
Db 118 PW-----NFVQ--CAKPEKQKEWTRTAKWKQNFKCNKEIYVVTESYKTLA 168
Qy 314 IRDLFGPLDQTPRIVILQGAAGIGKSTLQVKEAWGRGOLYGDGFQHFYFSCRELA 373
Db 169 ---LCNPKIETPFAHAILVHGGPGSGKTTMAKQMLEWSESK-QAQIFSCAFYISCREVN 224
Qy 374 QSKVSLAELIGKDGATPAPRIQLSRPERLLFILDGVDEPGWVLOEPSSSELCHWSQ 433
Db 225 NTKPCTFHLLSMDNPSWRDCVIRDILGKEFLFVVDGDELTFPAGALIRDLGDMNTV 284
Qy 434 QPADALLGLGKTLTPEASFLITARTALONLIPSLQARVVEVLGFSSSRKEYFYR 493
Db 285 KPVEVLGSLKKRMAPHATLTVTRTQSLHQIFVMMDQLLVETLGFLEQKEQYFQY 344
Qy 494 FTDER-----QAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKILTS 544
Db 345 FEDEEGEEDKGBKALRAKALKEVRCNADLYQMASLPTACGIFCLCLELRMKKGEDLSLTC 404
Qy 545 KTTTLCILHYLAQALQPLGQPLRD-----LCSLAAGIWKQKTLFSPDDLKKGHL 596
Db 405 QTYTSMFLNFCVFSSETCEDHNEFQLIFKKICILAANSLEQVPLICEEDF----- 459
Qy 597 DGAIISTFLKMGILQHPILSY-----SFHLCQEFPEAANSYVLE---DEK 641
Db 460 -----LTKLNLNNHPWVRHILFKDSSSTHCLSFICLGQQLAAIIFVQELGQESK 513
Qy 642 GRGKHS-NCIIDILEKTEAYGIHGLFGASTTRFLLGLLSDGEREMENIFHCRLS----- 695
```

```
Db 514 GVSQYIIONMLSRARLKNPDLGSL-----PFVGLLNTRIQLKTTFCQISTEVKR 568
Qy 696 -----QGRNLNQWVPSLQLLQPHLSLESLSHCLYETRNKTLFTQVMAHFEENGMCVETDM 749
Db 569 KPLCESEGEN-----KPLLLMMNQELISCLVSEGEFVKEAMVLFDISLHLKTS 621
Qy 750 ELLVCTFCIKFSRHVKQLQ-L-IE-----GQRHRSWTSPWVLFWRVWVVD 794
Db 622 DLIIHASFCLKNSQNLQTMSLKVERAVFPENVAALSTAKHQRSPDEQRL----- 671
Qy 795 AYQWILFSLVKTNRNLKELDLGNSLSHSAVKSCKTLRRPRCLLETLRACGLTAECD 854
Db 672 TFWTDFCDTFNSNKKLVFELDIHESPLNSSALEICEKLPSASCCLOKVVK--NISPPDA 729
Qy 855 -KDLAFGLRANOTLTDELDFSNVLTDAKAHLQRLROPSCCKLQRLQVSCGLTSDCCQD 913
Db 730 YEKULCIENGKYKTIHLLQGNL-DSMHSLCEVLKPNPACNLKFLSLGSCSTAQAQWDD 788
Qy 914 LASVLSASPSLKELDLQNNLDVGVRLLCRGLRHP 949
Db 789 FFPVLKVNQSLIFDLDTNSLLDKSAKLLCNWKEP 824

RESULT 11
Q9EPG7 PRELIMINARY; PRT; 657 AA.
AC Q9EPG7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 23, Last annotation update)
DE RNI-like protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=213110002; PubMed=11416212;
RA Lane R.P., Cutforth T., Young J., Athanasiou M., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT loci.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395 (2001).
DR EMBL: AF321233; AAG45188.1; -.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTFase.
DR PROSITE; PS50503; LRR_RI; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 657 AA; 75410 MW; F4BED9E4BA19AAF3 CRC64;

Query Match 9.0%; Score 676; DB 11; Length 657;
Best Local Similarity 27.1%; Pred. No. 4.8e-42;
Matches 194; Conservative 95; Mismatches 220; Indels 206; Gaps 11;

Qy 314 IRDLFGPLDQTPRIVILQGAAGIGKSTLQVKEAWGRGOLYGDGFQHFYFSCREL 372
Db 78 LQHI FDEIDQISEAQTVVLQGAAGIGKTLTKKAVLEWADGNLY-QQTHVFLYNGKEI 136
Qy 373 AQSKVVSIAELIGKDGATPAPRIQLSRPERLLFILDGVDEPGWVLOEPSSSELCHWSQ 432
Db 137 SQVKEKSPAQLISKHWPSSSEGIEQ-----GQVEKGRVEE----- 172
Qy 433 POPADALLGLGKTLTPEASFLITARTALONLIPSLQARVVEVLGFSSSRKEYFYR 492
Db 173 ----- 172
Qy 493 YFTDERQAIRFLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKTLTSTTTTLC 552
Db 173 -----TCTSTALFT 182
Qy 553 HYLAQALQAPLIG-----PQLRDLCSLAEGIWQKTLFSPDDLKHLGDLAISTPL 605
```

```
Db 183 YYICSLFPRIPVGCVTLPNETLRLSLCKAAVEGIWTKHVLVQONLRKHELTREDILFEL 242
Qy 606 KMGILQEH-PIPLSYSFHLCQEPFAAMSIVL-EDEKGRGHSNCIIIDLEKTLAYGTH 663
Db 243 DAKVLOQDTEYENCYMFTHLVQEPFAALFYLLRENLEEQDYPSEPFENLYLLESNHHT 302
Qy 664 GLFGASTTRFLGLLSDGEREMENIFHCRLSQ--GRNLMQWPSLQ---LLLOPHSLE 717
Db 303 DPHEQMKCFGLNKNDRVQLEETFNLTISNEVREELACLEGLEKDDSSLSQRFOD 362
Qy 718 SLHCLYETRNKTLTQMAHFEEMGCVETDMELLVCTFCIKFSRHVKXKLQIEGRQHS 777
Db 363 LLHCIVETQOEFTQALVFOKIIIVRDEEPQLRIYSFCLKCHTCLKMRLTARADLN 422
Qy 778 TWSPTMVLFRWPVTDAYQWILFSLVKVTRNLKELDLS----- 815
Db 423 MLDTAEMCLEGAQVVIHYWQDLFSVLHTNESLIEMDLYESLDESIMKILNEELSHPKC 482
Qy 816 -----SGNSLSH-----SAVSLCKTLRRPRCLLETL 842
Db 483 KLOKLFRSVDFLNGQDFTFLASNKKVTHLDLKETDGLVNGLKTLCCEALKCKGCKLRLV 542
Qy 843 RLACGLTAEDCKDLAFGRANQTLFELDLSFNVLTDAGAKHLCORLQPSCKLQRLQV 902
Db 543 RLASCDLNAVRCQKLSNALQTRSLVFLNLSNLSNDGVKSLCEVLENPNSSLSRLALM 602
Qy 903 SCGLTSDCCDLASVLSASPSLKELDLOQNLDVGVRLLCBGLRHPACKLRLG 957
Db 603 SCVLTSKACCDLASVLVNSNLSWLSLDGHNILCDAGLNLCDALRPNCHVQRLG 657

RESULT 12
Q8CCN1 PRELIMINARY; PRT; 673 AA.
ID Q8CCN1
AC Q8CCN1;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE PAAD and NACHT containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK032446; BAC27872.1; -.
SQ SEQUENCE 673 AA; 76367 MW; EE773CS92BEC7054 CRC64;

Query Match 7.6%; Score 573; DB 11; Length 673;
Best Local Similarity 33.9%; Pred. No. 2.9e-34;
Matches 152; Conservative 81; Mismatches 154; Indels 62; Gaps 14;

Qy 312 IBTDLFGGLDT-QBPRVILQGAAGIGKSTIARQKVAWGRGQLYGDRFOHVFVFSR 370
Db 146 VDVETLFAPEAESYSTPPIVWQGSAGTKTLVKLVQDWSKGKLYPCQFDYVVFVSCR 205
Qy 371 ELAQSKVSLAEIJ-----KGGTATPAPIRQILSRPERLLFILDGVDEPGWVLPQSPSEL 426
Db 206 EVLLPKCLDPLNICWCCGDD---QAPVTEILRQGRLLFILDGYDE---LQKSSR-- 255
Qy 427 CLHWSQPADALLGSLGKTLIPASFLITARTALONLISLEQARWVEVLGSESSR 486
Db 256 -----ABCVHITLMRRREVP-CSLDITRPPALQSLPEMLGRRHVLVLGFSSEER 305
Qy 487 KEYFYRYFTDERQAIRAFRLVSKNELWALCLVPWVSWLACTCLMQMKRKEKLTILTS 546
```

```
Db 306 ETYFSSCFTDEQLKNALEFVQNNAVLYKACQVPGICWVCSWLKKQARGQEVSETPSN 365
Qy 547 TTITLCHVLAQALQAPLGPO-----LRDLCSLAAGIWKQKTLFSPDDLRKHGLDG 598
Db 366 STDIFAVVSTFLPTDGNQDSSELTRHKVLSKSLCSLAAEGMRHORLLFEEVRLRKHGLDG 425
Qy 599 ATISTFLK-----MGILQEHPIPLSYSFHLCQEPFAAMSIVLEL---EKGRGHSN 648
Db 426 PSLTAFINCIDYRAGLGKK-----FYSFRHISFQEPFYAMSLVKEDSQOQGEATHKE 479
Qy 649 C--IIDLEKTLAYGIHGLFGASTTRFLGLLSDGEREMENIFHCRL--SGRNLMQW 704
Db 480 VAKLVDPNHEEV-----TSLQFLFDMKLTGKTLISGLKFCFRITAPSVRQDLKHFK 531
Qy 705 PSLQQLLQPHLSLESILHCLYETRNKTLFQ 733
Db 532 EQIEAIKYKRWLDLEFSLYDSKIKK-LTQ 559

RESULT 13
Q8IWF5 PRELIMINARY; PRT; 953 AA.
ID Q8IWF5
AC Q8IWF5;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Caspase recruitment domain family, member 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040339; AAH40339.1; -.
SQ SEQUENCE 953 AA; 107671 MW; 0A9DF167BE87E21A CRC64;

Query Match 5.2%; Score 393.5; DB 4; Length 953;
Best Local Similarity 24.0%; Pred. No. 1.8e-20;
Matches 214; Conservative 131; Mismatches 336; Indels 209; Gaps 38;

Qy 237 VVGTPQAHTSLQPHHPWEPVRESLCSLSTWPKNEDFNQKFTQLLLQRPSPQDPLV 296
Db 120 VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAKK--BELLLEIYMDTIMELV 167
Qy 297 KRSWPDYVVEENRGHLIEIRDLFG--PGLDQEPRIVLQGAAGIGKSTIARQKVAWGRG 354
Db 168 -----GFSNESLGSLSLACLDDHTTGILNEQGETIFILGDAGVGKSMQLQRLQSLMATG 222
Qy 355 QL-YGDRFQHVYFSCRELA---QSKVVSIAELIKGD---GTATPAPIRQILSR-PERLL 406
Db 223 RLDAGVKF--FFHFRCRMFSCKESDRLCQDLFLFKHYCYPERDPEEVFAFLRPPHAL 280
Qy 407 FILDGVDEPGWVLPQSPSELCLH-----WSQPPADALLGSLGKTLIPAEASFLITA 458
Db 281 FTFDGLD-----ELHSDLDLSRVDPSSCPWEPAHPL-VLLANLSKLKAGSKLITA 332
Qy 459 RTTALQNLIPSLQARWVEVLGFSSESSRKEYFYRYFTDERQAIRAFRLVSKNELWALCL 518
Db 333 RTGI---EVPRQFLRKKVLLRGFSFSLRAYARRMFPERALQDRLLSLEANPNLCSLCS 389
Qy 519 VPWVSWLACTCLMQ-----OMKRKEKLTILTSKTTTL 550
Db 390 VPLFCWIIIFRCFQHPRAAFEGSPQLPDCDTMTLTDVFLLVTEVHLNRMQPSSLVQENT--- 446
Qy 551 CLHYLAQALQAPLQPLQRLD-LCSL---AAGIMQKTLFSPDDLRKHGLDGAIIITFLK 606
Db 447 --HSPVETLHAG-----RDTLCSLGQVAHRGMKSLFVFTQEEVQASGLQ-----ERDMQ 494
Qy 607 MGILQEHF-----IPLSYSFHLCQEPFAAMSIVLEDEKGR-----GK 645
```

```

Db 495 LGFLRALPELPGGQOQSEFFHLLTQAFFTAFFLVDVRVGTQELLRRFFQEMPPAGAA 554
QY 646 HSNCI-----IDLEKTLAYGHTGLFGASTTRFFLLGLLSDEGE-----683
Db 555 TTSCYPPPLPFCQLOQSGPAREDLFKNKD-----HFQF-----TNLFLCGLLSKAKQKLLRHL 607
QY 684 -----REMENIFHCRSLQGRNLMQWVPSL-----QLLQPHSLSLHCLYETRNKTF 730
Db 608 VPAALRRKKALWAHLFSLRGYLSKSLPRVQVESFNQVQAMPTFIWMLRCIVETQ-----663
QY 731 LTQVMAHFEMGMCVETDMLLVCTFC-----IKFSRH--VKKLOI-----IEGR 773
Db 664 -SQVQQLAARGICANY-LKLTVCNACSADCSALSFLVHHPKRLALDLNNLNLDYGV 721
QY 774 QHRSTWSPWVLFVRVPTDQWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCKTLR 833
Db 722 ELQPCFSRLTVLRSVNIQTDGGVKVLSBELTKYKIVTYGLYNNQITDVGARYVTKILD 781
QY 834 RPRCLLETURLAGCLTAEDCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLQRLR-QP 892
Db 782 ECKGLTH-LKLGKNKITSEGGKYLALAVKNKSISEVGMWGNQVDEGAKAPAEALRNHP 840
QY 893 SKLQRLQLVSCGLTSDCCODLASVLSASPSLKELDLOQNNLDDVGRVLLCEGLR-HPAC 951
Db 841 S--LTTLSLANSIGISTEGGKSLARALQONTLSLEILWLTQNLNDEVAESLAEMLKVNQTL 898
QY 952 KLIRLGLDQTT-----LSDME-----QELRALEQEK 978
Db 899 KHLWLQNIQITAKGTQALADALQNSGTGITEICLNGNLKPEEAKVYEDEK 948

RESULT 14
Q8K3Z0 PRELIMINARY; PRT; 1020 AA.
AC Q8K3Z0;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE CARD15.
GN CARD15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Iwanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,
RA Baugh M.M., Suing C., Rosenbaum J.T.;
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15
RT gene";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF520774; AAM76073.1; -
DR MGD; MGI:2429397; Card15.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR001130; TATD_DNase.
DR PROSITE; PS50209; CARD; 2.
DR PROSITE; PS50503; LRR_RI; 1.
DR PROSITE; PS50837; NACHT; 1.
DR PROSITE; PS01137; TATD; 1.
DR PROSITE; PS01137; TATD; 1.
SQ SEQUENCE 1020 AA; 113561 MW; 25504905ECF70FBB CRC64;

Query Match 5.2%; Score 392; DB 11; Length 1020;
Best Local Similarity 25.2%; Pred. No. 2.5e-20;
Matches 186; Conservative 122; Mismatches 316; Indels 114; Gaps 28;

QY 311 LIIEIRDLFGP-GLDTQBPRIIVILQAGAGICKSTLARQVKEAGRGQLYGDRFQHVYFSC 369
Db 255 ILGLEDLFDTHGHLNRDADTILVVGAGSGKSTLQRLHLLWATGRSF-QEFLFIFFPFC 313

370 REL-AQSKVSVSLAEILGK-----DGTATPAPIRQILSRPERLLFILDGVDPGWVLOEPSS 424
Db 314 ROLQCVAKPLSLRTLLFHHCCWPDVAQDDVFOFLDHPDRVLLTDFGLDEFKFRFTDRR 373
QY 425 ELCLHWS--QPQPADALLGSLGKLTLPASFLITARTTALQNLIPSLDQARWVEVLGFS 482
Db 374 ----HCSPIDPTSVQTLFNLQGLLKNACKVLTSPDVAVALLRKKFVTE-LQKLGFS 428
QY 483 ESSRKEYFYRYPTDRQAIRAFRLVKSKELWALCLVPWVSWLACTCLMQOKRKEKLT 542
Db 429 EEGIQLYLRKHREPCGVADRLIQLIQATSAHLGLCHLPVFSNMVSRCHRELLQNRGFFT 488
QY 543 TSKTTTTLCL-HYLAQALQAP-----LGP-----QLRDLCSLAAGIWKQKTLF 586
Db 489 TSDMYLLIQLHFL---LHASPPDSPLGLGPGLLQSRSLTLLHGLHALRGLAMSCYVF 545
QY 587 SPDDLRLKGLDCAIISTFLKMGIL--QEHPIPLS---YSFIHLCOEERFAAMSYYVLEDEK 641
Db 546 SAQQLQAAQVSDDIS---LGLFVRAQSSVPGSKAPLEFLHITFOCFFAAFVLAVSADT 601
QY 642 -----GRGKHNSCI--IDLEKTLAYGI-----HGLFGASTTRFFLL 675
Db 602 SVASLKLHLPFCGLGSSLLGRLLPNLCTIOGSRVKGSEAAALQKAPHNL--QITAFLA 659
QY 676 GLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSL-----ESLH 720
Db 660 GLLS---QQHRLLAACQVSE-RVLLQROARARSCL-AHSREHFHSIPPAVPGETKSMH 714
QY 721 C-----LYETRNKTLTQVMAHFE-----EMGMCVETDMLLVCTFCIKFSRHVK 766
Db 715 AMPGFTWLIRSLYEQEQLAQEAVERRLDIGHLKLTCFVRGPAECAALAFVLQHLQRPVA 774
QY 767 LQL-----IEGRQHRSTWSPWVLFVRVPTDQWQILFSLVKVTRNLKELDLSGNS 819
Db 775 LQLDYNSVDGVGEQLRPLCGVCTALYLRDNNISDEGARTLVECALRCEQLKALFNKK 834
QY 820 LSHSAVSKLCKTLRRPRCILETLRLAGCLTAEDCKDLAFGLRANOTLTLDLSFNVLTD 879
Db 835 LTDACACSMAKLLAHKQNEL-SLRVGNHNTAAGAEVLAQGLKSNSTSLKFLGFWGNSVD 893
QY 880 AGAKHLQRLRQPSCKLQRLQLVSCGLTSDCCODLASVLSASPSLKELDLOQNNLDDGV 939
Db 894 KGTQALAEVYADHQ-NLKWLSLVGNNGSMGAEALALMLEKNKSLBELCLEENHICDEGV 952
QY 940 RLLCEGL-RHPACKLIRL 956
Db 953 YSLAEGLKXNSTLKFLKL 970

RESULT 15
Q9BY26 PRELIMINARY; PRT; 287 AA.
AC Q9BY26;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Leucine-rich-repeat protein RNO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shami P.J., Kanai N., Wang L.Y., Vreeke T.M., Parker C.J.;
RT "Identification and characterization of a novel gene that is
RT upregulated in leukemia cells by nitric oxide.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231021; AAK14942.1; -
DR HSSP; P13489; 1A4Y.
DR InterPro; IPR007091; LRR_RNinh.
DR PROSITE; PS50503; LRR_RI; 2.
SQ SEQUENCE 287 AA; 31765 MW; BD3816C3255B2F9E CRC64;
```

Query Match 5.1%; Score 383; DB 4; Length 287;
Best Local Similarity 43.7%; Pred. NO. 1.7e-20;
Matches 93; Conservative 24; Mismatches 94; Indels 2; Gaps 1;

Qy	769	LIEGROHRSWPTMVLFRWVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKS	828
Db	73	LCBGLRHPQ--CRLQMIQLRKCOLESQACQEMASVLGTNPPLVVELDTGNALEDLGLRL	130
Qy	829	CKTLRRPRCLLETRLAGCGLTAECDKDLAFGLRANQTLTELDLSFNVLTDAGAKHL	888
Db	131	CQGLRHPVCELRFLWKICELTAAACDELASTLSVNSQSLRELDLSNELGDLGVLL	190
Qy	889	LROPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRL	948
Db	191	LRHPTCKLQTLRLDSCGLTAKACENLVFTLGINQTLTDLVLTNNALGDTGVRL	250
Qy	949	PACKLIRLGLDQTTLSDEMROELRALEQEKPOL	981
Db	251	PGCKLRVLMFGMDLNKMTSHSLAALRVTKPYL	283

Search completed: January 29, 2004, 13:48:11
Job time : 84.1934 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:11:31 ; Search time 28.1576 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWRLACYLEFLKKEE.....HLIMELWEKSKKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	388.5	5.2	953	4	US-09-999-041A-8
2	388.5	5.2	953	4	US-09-245-281-8
3	388.5	5.2	953	4	US-09-207-359B-8
4	388.5	5.2	953	4	US-09-340-620A-8
5	363.5	4.8	456	2	US-08-910-731-8
6	361.5	4.8	456	2	US-08-910-731-4
7	361.5	4.8	456	2	US-08-795-395-4
8	357	4.7	461	2	US-08-910-731-6
9	354.5	4.7	456	2	US-08-910-731-2
10	354.5	4.7	456	2	US-08-795-395-2
11	348	4.6	953	4	US-09-245-281-43
12	348	4.6	953	4	US-09-207-359B-43
13	348	4.6	953	4	US-09-340-620A-43
14	345	4.6	1130	2	US-08-519-547A-6
15	326.5	4.3	966	4	US-09-207-359B-47
16	253.5	3.4	193	4	US-09-340-620A-61
17	250.5	3.3	195	4	US-09-340-620A-49
18	218.5	2.9	490	4	US-08-999-041A-26
19	218.5	2.9	490	4	US-09-245-281-26
20	218.5	2.9	490	4	US-09-207-359B-26
21	218.5	2.9	490	4	US-09-340-620A-26
22	216	2.9	71	4	US-09-340-620A-58
23	209	2.8	70	4	US-09-340-620A-57
24	209	2.8	70	4	US-09-340-620A-66
25	198	2.6	1151	3	US-08-836-134-23
26	198	2.6	1151	4	US-09-493-784-23
27	198	2.6	1232	3	US-08-836-134-2

28	198	2.6	1232	4	US-09-493-784-2	Sequence 2, Appli
29	179.5	2.4	483	4	US-09-904-615-154	Sequence 154, App
30	175.5	2.3	200	4	US-09-099-041A-11	Sequence 11, Appl
31	175.5	2.3	200	4	US-09-245-281-11	Sequence 11, Appl
32	175.5	2.3	200	4	US-09-207-359B-11	Sequence 11, Appl
33	175.5	2.3	200	4	US-09-340-620A-11	Sequence 11, Appl
34	165	2.2	1466	4	US-09-252-991A-30085	Sequence 30085, A
35	147.5	2.0	793	3	US-09-012-710-10	Sequence 10, Appl
36	147.5	2.0	793	4	US-09-556-273-10	Sequence 10, Appl
37	147	2.0	2482	1	US-08-328-254-6	Sequence 6, Appli
38	143.5	1.9	4302	4	US-09-052-469-8	Sequence 8, Appli
39	143.5	1.9	4302	4	US-08-422-582-8	Sequence 8, Appli
40	143.5	1.9	4339	4	US-09-052-469-6	Sequence 6, Appli
41	143.5	1.9	4339	4	US-08-422-582-6	Sequence 6, Appli
42	142.5	1.9	4302	3	US-08-658-136-5	Sequence 5, Appli
43	142.5	1.9	4303	2	US-08-460-751-2	Sequence 2, Appli
44	138.5	1.8	794	1	US-08-393-333-2	Sequence 2, Appli
45	138.5	1.8	794	3	US-09-087-465-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-099-041A-8
; Sequence 8, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-8

Query Match		5.2%	Score 388.5;	DB 4;	Length 953;
Best Local Similarity		24.6%	Pred No. 3.7e-28;		
Matches		217;	Conservative 129;	Mismatches 343;	Indels 193; Gaps 39;
QY	237	VVGTPPQAHTSLQPHHPWPSVRESLCSTWPKNEDFNQFTQLLLQRPSPQDPLV	296		
DB	120	VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAKK-EELLLEEIVMDTLMELV	167		
QY	297	KRSWPDYVENRCHLIEIRDLFG--PGLDTQPRVILQGAAGIGKSTLARQVKEAWGRG	354		
DB	168	-----GFSNESIGSLNSLACLLDHTTGILNEOGETIFILGDAGVGKSMQLQLQSLWATG	222		
QY	355	QL-YGDRFQHFVYFSCRELA---QSKVWSLABLIGKD---GTATPAPIQILSR-PERLL	406		
DB	223	RLDAGVKF--FFHRCRMFSCFKESORICLQQLLFKHYCPERDPEVFAFLRPFHVAL	280		
QY	407	FILDGVDFEGWVLOQFSSSELCLH-----WSQPQADALLGSLGLTKILPEASFLLTA	458		
DB	281	FTFDGLD-----ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLGASKULLTA	332		
QY	459	RTTALQNLIPSLQARWVEVLGFSRSSRKEYFYRFTDERQAIRAPRLVKSNNKELWALCL	518		
DB	333	RTGI---EVPRQFLRKKVLLRGFSFSLRAYARRMFPERALQDRLLSLEAENPNCSLCS	389		
QY	519	VPVWSWLACTCLMQQMKKE-----KLTJTSKTTTTLCLHY-----LAQALQAOA	563		
DB	390	VELFCWIIIFRCQHFRAAFEGSPQLPDCDTMTITDVFLLVTEVHLNMQSSVLQVRNTRSP	449		
QY	564	L-----GPQLRD-LCSL-----AAEGIWQKTLFSPDRLRKHGLDGAISTFLKMWILQHP	614		

Db 450 VETLHAG---RDTLCSLQVAHGMKSLFVFTQEEVQASGLQ-----ERDMQLGFLRALP 502
 Qy 615 -----IPLSYFHLFCFQEFFAAMSYVLEDEKGR-----GKHSNCI--- 650
 Db 503 ELPGGDQOQSEYFFHLTLQAFFTAFLVLDLDRVGTQELLRFFQEMMPAAGATTSCYPFF 562
 Qy 651 -----IDLEKLEAVGIHGLFGASTTFLGLLSDGB-----R 684
 Db 563 LPPQCLQSGGPAREDLFKNKD-----HFQF---TNLFLCGLLSKAKQKLLRHLVPAALRR 615
 Qy 685 EMENIFHRLSOGRLNMQWVPSL-----QLLQPHSLSHCLYETRNKTLFTQMAHF 738
 Db 616 KRKALWAHLFSSRLGYLKSILPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SOKVGOL 670
 Qy 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----IEGROHSTWSP 781
 Db 671 AARGICANY-LKLTVCNACSADCSALSFVLHHPKRLALDLDNNLNNDYGVRELOQCFSR 729
 Qy 782 TMVLFRRVPTDAYWQILFVSLKVTNKLKELDSLNSLSHSAVKSCLKTLRPRCLLET 841
 Db 730 LTVLRUSVNOITDGGVKVISELTXYKIVTYLGLYNNQITDVGARYVTTKILDECKGLTH- 788
 Qy 842 LRLAGCGLTAEDCKDLAFGLRANQTLTDLSPNVLTDAKAKHLQORL-OPSKCLQRLQ 900
 Db 789 LKLGKKNITSEGGKYALAVKNSKSISEVGMGNQVQDEGAKAFALRNHPS--LTTLS 846
 Qy 901 LVSCGLTSDCCQDLASVLSASPSLKDLDLQNNLDVGVRLLCGLR-HPACKLIRGLD 959
 Db 847 LASNGISTEGGKSLARALQONTSLLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLION 906
 Qy 960 QTT-----LSDEMR-----QELRALEQEK 978
 Db 907 QITAKGTAQLADALQONTGITEICLNGNLKPEEAKVYEDEK 948

RESULT 2

US-09-245-281-8
 ; Sequence 8, Application US/09245281
 ; Patent No. 6369196
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
 ; FILE REFERENCE: 07334/118001
 ; CURRENT APPLICATION NUMBER: US/09/245,281
 ; EARLIER FILING DATE: 1999-02-05
 ; EARLIER FILING DATE: 1998-12-08
 ; EARLIER FILING DATE: 1998-06-17
 ; EARLIER FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 953
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-245-281-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
 Best Local Similarity 24.6%; Pred. No. 3.7e-28;
 Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
 Qy 237 VVGTTPQAHTSLQPHHPWEPVSRESLCSSTWPKNEDFNQKFTQLLLQRPSPRSDIIV 296
 Db 120 VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAKK--EELLSEIYMDTIMELV 167
 Qy 297 KRSWPDYVENRGLHIEIDFLG--PGLDTPRIVILQAGIGKSTLARQKEAWNGR 354
 Db 168 -----GFSNESLSLSLACLHTTGTILNEQGETIFILGDAGVGKSMLLORLSLWATG 222

Qy 355 QL-YGDREHQHFVYFSCRELA---OSKVVSLAELIGKD---GTATPAPIROIILSR-PERLL 406
 Db 223 RLDAGVKF--FFHFCRMFSCFKESDRCLQDLFLFKHYCYPEDDEEVFAFLRPPHVAL 280
 Qy 407 FILDQVDFPGWVLPSPSELCLH-----WSQDQADALLGSLGLTKTILPEASFLITA 458
 Db 281 FTFDGLD-----ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKJLTA 332
 Qy 459 RTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYFTDERQAIRAFRLVKNKELWALCL 518
 Db 333 RTGI---EVRQFURKKVLLRGFSPSHURARRMFFERALQDRLLSQUEANPNLCSLCS 389
 Qy 519 VPVSWLACTCLMQOMKKE-----KLTLSTKTTTTLCLHY-----LAQALQAP 563
 Db 390 VPLFCWIIIFRCFQHPRAAFEGSPQLPDCMTWILTDVFLVTEVHLNRMQPSLSVQNRTRSP 449
 Qy 564 L-----GPOLRD-LCSL---AABGIWQKTTLPSPDDLRKGLDGAIIITFFKMGILQHP 614
 Db 450 VETLHAG---RDTLCSLQVAHGMKSLFVFTQEEVQASGLQ-----ERDMQLGFLRALP 502
 Qy 615 -----IPLSYFHLFCFQEFFAAMSYVLEDEKGR-----GKHSNCI--- 650
 Db 503 ELPGGDQOQSEYFFHLTLQAFFTAFLVLDLDRVGTQELLRFFQEMMPAAGATTSCYPFF 562
 Qy 651 -----IDLEKLEAVGIHGLFGASTTFLGLLSDGB-----R 684
 Db 563 LPPQCLQSGGPAREDLFKNKD-----HFQF---TNLFLCGLLSKAKQKLLRHLVPAALRR 615
 Qy 685 EMENIFHRLSOGRLNMQWVPSL-----QLLQPHSLSHCLYETRNKTLFTQMAHF 738
 Db 616 KRKALWAHLFSSRLGYLKSILPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SOKVGOL 670
 Qy 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----IEGROHSTWSP 781
 Db 671 AARGICANY-LKLTVCNACSADCSALSFVLHHPKRLALDLDNNLNNDYGVRELOQCFSR 729
 Qy 782 TMVLFRRVPTDAYWQILFVSLKVTNKLKELDSLNSLSHSAVKSCLKTLRPRCLLET 841
 Db 730 LTVLRUSVNOITDGGVKVISELTXYKIVTYLGLYNNQITDVGARYVTTKILDECKGLTH- 788
 Qy 842 LRLAGCGLTAEDCKDLAFGLRANQTLTDLSPNVLTDAKAKHLQORL-OPSKCLQRLQ 900
 Db 789 LKLGKKNITSEGGKYALAVKNSKSISEVGMGNQVQDEGAKAFALRNHPS--LTTLS 846
 Qy 901 LVSCGLTSDCCQDLASVLSASPSLKDLDLQNNLDVGVRLLCGLR-HPACKLIRGLD 959
 Db 847 LASNGISTEGGKSLARALQONTSLLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLION 906
 Qy 960 QTT-----LSDEMR-----QELRALEQEK 978
 Db 907 QITAKGTAQLADALQONTGITEICLNGNLKPEEAKVYEDEK 948

RESULT 3

US-09-207-359B-8
 ; Sequence 8, Application US/09207359B
 ; Patent No. 6469140
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-112001
 ; CURRENT APPLICATION NUMBER: US/09/207,359B
 ; CURRENT FILING DATE: 1998-12-08
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 953
 ; TYPE: PRT

i: ORGANISM: Homo sapiens
 US-09-207-359B-B

	Query Match	5.2%	Score 388.5	DB 4	Length 953
	Best Local Similarity	24.6%	Pred. No. 3.7e-28		
	Matches 217	Conservative 129	Mismatches 343	Indels 193	Gaps 39

QY	237	VVGTPPOAHTSLQPHHHHPWSPRESLCSTWPKNEDFNOKFTCLLLQRPHPSPQDPLV	296
DB	120	VVNTDPSVRSYQOURLHLGRDS-KFVLUC-----YAKC--EELLEIEYMDTIMELV	167
QY	297	KRSPWDVVEENRGHLIEIRDLFG--PGLDQEPRIVLQGAAGIGKSTLARQVKEAAGR	354
DB	168	-----GFSNESLSLNLACLLDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATG	222
QY	355	QL-YGRDPQHVFFVSCRELA---QSKVVSLAELIGKO---GTATPAPIQLLSR-PERLL	406
DB	223	RLDAGVKF--FFHFRCMFSCFKESDRLCQDLLFKHYCYPERDPEVPFAFLFPFHAL	280
QY	407	FILDGVDEPCWGLQEPSSELCH-----WSQPQPADALLGSLGKLTILPEASFLITA	458
DB	281	FTFDGLD-----ELHSDLDLSEVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKLLTA	332
QY	459	RTTALQNLIPSLBQARWVEVLGSESRKSYFYRYFTDERQAIARFLVKSXKELWALCL	518
DB	333	RTGI--EVPRQFLRKVKLLRGFSPLRAYARMFPERALQDRLLSQLEANPNLCSLCS	389
QY	519	VPVWSLACTCLMQOMRKE-----KLTLSKTTTTLCLHY-----LAQALQAP	563
DB	390	VPILFCWTFIRFOHFRAAFEGSPOLPCTMTLTDFLLVTEVHLNRWQPSLVORNRTRSP	449
QY	564	L-----GPQLRD-LCSLI---AABGIMQKTKLFPSPDLRKHGLDGAIISTFLKMILOEHP	614
DB	450	VETLHAG--RDTLCSLGQVAHRGMEKSLFVFTQEEVQASGLQ---ERDWQLGFLRALP	502
QY	615	-----IPLSYSFTHLCFQCFEFAAMSVYDEKGR-----GKHSNCI----	650
DB	503	ELGPGSQQSYEFFHLLTLOAFTFAFFVLDDRVGTQBLRFPFQEWMPPPAGAATTSCYPFF	562
QY	651	-----IDLEKTLAEGIHGLFGASTTRFLGLLSDEGE-----R	684
DB	563	LPQCLOGSGPAREDLPNKD-----HQFQ---TNLFLCGLLSKAKQKLLRHLVPAALRR	615
QY	685	EMENIFHCRLSQGRNLMQWVPSL-----QLLQPHSLESILCHLIYETRNTKTLFQVMAHF	738
DB	616	KRKALWAHLFSSLRGYKLSLPRVQVESFNQVQAMPTFTIMWLRCIYETQ-----SQKVQGL	670
QY	739	EEMCMCVETDMEILVCTFC-----IKFSRH--VKKLQL-----IBGQHRSTWSP	781
DB	671	AARGICANY-LKLTVCNACSDACSLSFVLHFFPRRLALDDNNLNLDYGVREIQQCFCSR	729
QY	782	TMVVLFRWPVPTDAYMOTILFSLVKVTNRNLKBLDLSGNSLSHSAVKSLCKTLRRPRCLLET	841
DB	730	LTVLRLSVNQITDGGVKVLSBELTKYKIVTVLYGLYNNQITDVGARYVTKILDECKGLTH-	788
QY	842	LRLAGCGLTABDCKDLAFGLRANOTLITELDLSFNVLTDAGAKHLCQRLR-QPSCKLQRLQ	900
DB	789	LKLGNKITSEGGKYLAALAVKNSKSI SEVGMWGNQVDEGAKAFEAALRNHPS--LTTLS	846
QY	901	LVSGLTSDCCODILASVLSASPSLKELDLQONNLDVGVRLCLCGLR-HPACKLIRLGLD	959
DB	847	LASNGISTEGGKSARALAQNTSLSEIWLWTQNELNDEVAESLAEMLVKNQTLKHLWLION	906
QY	960	QTT-----LSDENR-----QELRALEQEK	978
DB	907	QITAKGTAOLADALQOSTGITEICLNGNLIKPEEAKVYDEBK	948

RESULT 4
US-09-340-620A-8
; Sequence 8, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:

```

: APPLICANT: Bertin, John
:
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
:
: FILE REFERENCE: 07334-124001
:
: CURRENT APPLICATION NUMBER: US/09/340,620A
:
: CURRENT FILING DATE: 1999-06-28
:
: PRIOR APPLICATION NUMBER: US 09/245,281
:
: PRIOR FILING DATE: 1999-02-05
:
: PRIOR APPLICATION NUMBER: US 09/207,359
:
: PRIOR FILING DATE: 1998-12-08
:
: PRIOR APPLICATION NUMBER: US 09/099,041
:
: PRIOR FILING DATE: 1998-06-17
:
: PRIOR APPLICATION NUMBER: US 09/019,942
:
: PRIOR FILING DATE: 1998-02-06
:
: NUMBER OF SEQ ID NOS: 71
:
: SOFTWARE: FastSEQ for Windows Version 4.0
:
: SEQ ID NO 8
:
: LENGTH: 953
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: OS-09-340-620A-8

```

Query Match	5.2%	Score 388.5	DB 4	Length 953
Best Local Similarity	24.6%	Pred. No. 3.7e-28		
Matches	217	Conservative 129	Mismatches 343	Indels 193
Gaps	39			
Qy	237	VVGTPQAHTSLQPHHPWEPSPVRESICSTWPKWNEFNQKFTQLLQPHRPSQDPLV	296	
Db	120	VNMTDVSRYTOQLRHHLGRDS--KFVLC-----YAKQ--EELLLEIYMDTIMELV	167	
Qy	297	KRSWPDYVENRGHLEIRDLFG--PGLDTPQEPRIVILQCAAGIGKSTLARQVKEAWGRG	354	
Db	168	-----GFSNESLGSLSACLDDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATG	222	
Qy	355	QL-YGRDRFOHVFPVFSRELA---OSKVSJLAEILGKD---GTATPAPIQILSR--PERLL	406	
Db	223	RLDAGVKF--FFHRCRMFSCFKESDRCLQIDLLFKHYCYPERDPBEVFAFLRFPFHAL	280	
Qy	407	FILDGVDEPGWLVQEPSSSELCLH-----WSQOPADALLGSLGKKTILPEASFLITA	458	
Db	281	FTPDGLD-----ELHSDDLSRVPDSCPWPAPHL-VLLANLISGKLLKGASKLLTA	332	
Qy	459	RTTALQNLIPSLBEQARWEVLGVPSESRKEYFYRYFTDERQAIRAFRLVKSXKELMALCL	518	
Db	333	RTGI---EVPRQLRKVLLRGFSPSHLRAYARMPERALQDRLLSQLEANPNLCSLCS	389	
Qy	519	VPVSWLACTCLMQQMKRKE-----KLTLSKTTTTLCLHY-----LAQLAQAP	563	
Db	390	VPLFCWIIIFRCQHFRAAFEGSPQLDCTMTLTDVLLVTEVHLNRMQPSSLVQRNTRSP	449	
Qy	564	L-----GPQLRD-LCSJ---AAGIWKTKLTFSPDDLKGHLGDAIISTPLKWLGLQHP	614	
Db	450	VELTHAG--RDTLCSLGQVAHRGMBKSJVFVTOEEVQASGLQ---ERDMQJGFURALP	502	
Qy	615	-----IPLSYSFIHLCFQBFFAAMSJVLEDEKGR-----GKHSNCI---	650	
Db	503	ELGPGGQQSYEFHLLTQAFFTAFLVLDDRVGTQELLRFFQEWMPPGAATTCVPPF	562	
Qy	651	-----IDLEKTEAYGIHGLFGAATTRFLLGLLSDEGE-----R	684	
Db	563	LPFQCLQGSGPAREDLPKND---HFQF---TNLFLCGLLSKAKQKLLRHLVPAAALRR	615	
Qy	685	EMENI FHCRLSOGNLMQWVPSL-----QLLLQPHSLSLSHLCYETRNKTKTFLTVMAHF	738	
Db	616	KRKALWAHLFSSLRGYKLSLPRVQVSEFNGVQAMPTFIWMLRCIYETQ-----SQRVQGL	670	
Qy	739	EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----IEGRQHRSTWSP	781	
Db	671	AARGI CANY--LKLTYCNACSDACSLSFVLHHPFKRLALDDNNLNDYGVRELQPCFSR	729	
Qy	782	TMVVLFRWPEVTDAYQWILFSLVKVTRNLKELDLSGNSLSHSAVKSICKTLRPRCLLET	841	
Db	730	LTVLRLSVNOITDGGVKVLSSEELTKYKIVYGLGYLXNOITDVGARYVTKLIDBCKGLTH--	788	

RESULT 4

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-731-4

Query Match 4.8%; Score 361.5; DB 2; Length 456;
Best Local Similarity 35.5%; Pred. No. 4.2e-26;
Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;

QY 792 VTDAYWQILFSLVKTRNLKELDLGNSLSHSAVKSLCKTLRPRCLLTETLRAGCGLTA 851
DB 206 ITSANKDCLDVVASKASIQELDLGNSKNGTGAALCSGLLPSRLRTLWLDVCDVTA 265
QY 852 EDCKLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTS--- 908
DB 266 EGCKDLRVLRKQSLKELSLAGNELKDEGAQLLCESLLEPGQLESLSWKTSLSLAASC 325
QY 909 -----DC-----CQDL 914
DB 326 PHFCSVLTKNSSLFELQMSNPLGDSGVVELCKALGYPTVLRVWLWGDCDVTDSGCSL 385
QY 915 ASVLSASPSLKELDLQNNLDDVGYRLLCBGLRHPACKLIRGLDQTLTSLDEMROELRAL 974
DB 386 ATVLLANRSLRELDLSNCGMDNGVLQLESKQPSCLLQQLVLYDIYWTDEVEDQRLAL 445
QY 975 EQEKPQLLIFS 985
DB 446 EEPSPSLRIIS 456

RESULT 7

US-08-795-395-4
Sequence 4, Application US/08795395
Patent No. 5965399

GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Cloning and Expression of Rat Liver and
TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,395
FILING DATE: 04-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440002

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-395-4

Query Match 4.8%; Score 361.5; DB 2; Length 456;
Best Local Similarity 35.5%; Pred. No. 4.2e-26;
Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;

QY 792 VTDAYWQILFSLVKTRNLKELDLGNSLSHSAVKSLCKTLRPRCLLTETLRAGCGLTA 851
DB 206 ITSANKDCLDVVASKASIQELDLGNSKNGTGAALCSGLLPSRLRTLWLDVCDVTA 265
QY 852 EDCKLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTS--- 908
DB 266 EGCKDLRVLRKQSLKELSLAGNELKDEGAQLLCESLLEPGQLESLSWKTSLSLAASC 325
QY 909 -----DC-----CQDL 914
DB 326 PHFCSVLTKNSSLFELQMSNPLGDSGVVELCKALGYPTVLRVWLWGDCDVTDSGCSL 385
QY 915 ASVLSASPSLKELDLQNNLDDVGYRLLCBGLRHPACKLIRGLDQTLTSLDEMROELRAL 974
DB 386 ATVLLANRSLRELDLSNCGMDNGVLQLESKQPSCLLQQLVLYDIYWTDEVEDQRLAL 445
QY 975 EQEKPQLLIFS 985
DB 446 EEPSPSLRIIS 456

RESULT 8

US-08-910-731-6
Sequence 6, Application US/08910731
Patent No. 5932440

GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
APPLICANT: SHANDILYA, HARINI
TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

```

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-910-731-6

Query Match      4.7%; Score 357; DB 2; Length 461;
Best Local Similarity 37.9%; Pred. No. 1.2e-25;
Matches 86; Conservative 30; Mismatches 91; Indels 20; Gaps 3;

QY 768 QLTEGQHRSTWSPVWVLFPRWVPTDAYWQI-----LFSVLKVTNRNLKELDSGN 818
DB 246 ELCPLGLHSSRLRTU-----WEGITAKGGCDLCRVLRAKESKELSLAGN 294
QY 819 SLSHSAVSKLCTLRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQTTLELDSFNVL 878
DB 295 ELGDEGARLLCETLLEPGQQLSLWKVSCSFTAACCPHFSSVLAQNRFLLEQLQISNNRLE 354
QY 879 DAGAKHLQRLRPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVG 938
DB 355 DAGVRELCOGLQPGSVLRVWLADCDVSDSCSLAATLLANHSRLRELDLSNNCLGDAG 414
QY 939 VRLLCGLRHPACKLRIGLDQTTLSDEMREQLRALEQEKPOLLIIPS 985
DB 415 ILQVESVQPGCLLEQLVLDYIWSEEMEDRLQALEKPKSLRVS 461

RESULT 9
US-08-910-731-2
; Sequence 2, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
US-08-910-731-2

Query Match      4.7%; Score 354.5; DB 2; Length 456;
Best Local Similarity 34.3%; Pred. No. 2e-25;
Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

QY 752 LVCTCFIVSFHVKKQLQIEGRHSTWSPVWVLFPRWVPTDAYWQIQLFSLVKVTRNLK 811
DB 127 LLCEGLLDPQCCHLEKQL-----EYCRLLTAASCEPLASVLRATRAK 168
QY 812 ELDLGNSLSHSAVSKLCTLRPRCLLETLRLAGCGLTAEDCKDLAFGLRANQ-TLTEL 870
DB 169 EUTVSNNDIAGEARVLGGGLADSACQLETLRLENCGLTTPANCKDLC-GIVASQASLREL 227
QY 871 DLFSFVLTDAGAKHLQRLRPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQ 930
DB 228 DLGNSGLDAGIAELCPGLSPASRLKTLWLWECDDITASGCRDLQVLAQKELSLA 287
QY 931 QNNLDDVGRLLCEGLRHPACK-----LIRLGLDQTT 962
DB 288 GNKLGDEGARLICESLLOPGQQLSLWKVSCSLTAACQHVSLMLTONKHLLEQLSSNK 347
QY 963 LSDENRQEL-RALEQEKQLLIFFSRKPSVMTPTTEGLDTGEMSNS-TSSLKQRLGSERA 1020
DB 348 LGDSGIQELCQALSQFGTTLRVLC-----LGDCEVINSGCCSL-----A 386

RESULT 10
US-08-795-395-2
; Sequence 2, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
US-08-910-731-2
```

```

RESULT 12
US-09-207-359B-43
; Sequence 43, Application US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; CURRENT FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 43
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-207-359B-43

Query Match 4.6%; Score 348; DB 4; Length 953;
Best Local Similarity 21.6%; Pred. No. 3.4e-24;
Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

QY 234 WAAVGTTPQAHTSLQPHHPWPSVRESLCTWPKWNEFQKFTQLLLQRP-----287
DB 103 WLSEIGFSP-----SOLIRTKTIIVTDPVSRYTQQLRHQGRDSKFM 144
QY 288 --HPSQDPLVKRSPD-----YVENRGHLEIRDLF--GPGDLOEPRIVILQGA 336
DB 145 LCYAKEDLLLEETVMDTLMGLVGFNNENLGLDCLDHDSTGVINEHGETVFVFGDA 204
QY 337 GIGKSTLARQVKEAWRGQLYGDRFQHVYFSCRELA---QSKVSLAEELIGKD---GTA 390
DB 205 GVGKSMLLQRLQSLWASGRLTSTA-KFFHFRCRMFSCFKESDMLSLQDLLFKHFCYCEPQ 263
QY 391 TPAPTRQILSR-PERLLFILDGVD-----PGWVLOEPPSELCLHWSQPOPADALLGSLGK 446
DB 264 DPEEVFSFLRPHTALTFTDGLDELHSDFDLSRPVDS--CCPWEPAHPL-VLLANLISG 320
QY 447 TILPEASFLITARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYTDERQAIAPRL 506
DB 321 RLLKGAGKLLTARTGV---EVPRLRKKVLLRGFSPSHLRAYARMPPERTAQEHLLQ 377
QY 507 VSKNELMALCLVPWVSWLACTCLMQ-----SOLIRTKTIIVTDPVSRYTQQLRHQGRDSKFM 144
DB 378 LDANPNLCSLCGVPLFCWIIIFRCFOHQTVPFEGSSQLPDCAVTLTDVFLVTEVHLNRP 437
QY 538 EKLTLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWQKTLFSPDRLKRGHD 597
DB 438 QPSSLVQNRTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEVQASKLO 490
QY 598 GAIISFLKMGILQEHPP-----IPLSYSFHLCFOEFAAMSYLEDE-----640
DB 491 ---EGDLQGLFRALPDVGPQGGQYEFPHLTLOAFTAFPLVADDDKYSTRELLRFFRE 546
QY 641 ---KRGKHSNCIIDLKTLAYGHLGFGAS-----VPSL-----QQLLOPHSLESILHCLY 723
DB 547 WTSPEGATSSSCHSP-----FSPQCLGSRSLGPDPRNKHDFQFTNLFVCGLLAKAR 600
QY 683 EREMENIFHCRLSOGRLNMQW-----VPSL-----QQLLOPHSLESILHCLY 723
DB 601 QKLLRLQVPKAILRRKRKALWAHLFASLSRYLSKSLPRVQSGGFNQVHAMPTFLWMLRCIY 660
QY 724 ETRNKTFLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHKVKKLQLEGRQHRSTWSP 783
DB 661 ETQ-----SQKVGRLAARGI-----SDYLUKLAFCNACSADCSALSFVLHFRQL-----706
QY 784 VVLFRRVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSVKSCKTLRRPRCLLETLR 843
DB 707 -----ALDDNNLNNDYGVQELQPCFSR-----LTIVR 734
QY 844 LACCGLTABDCXDLAPGLRANQTLTBLDSFNVLTDAGAKHLCQRLQPSCK-LQRLQV 902
DB 735 LSVNQITDGVKVLCELTAKYIVTFLGYNQITDIGARYVAQILDE--CRGLKHLKLG 792
QY 903 SCGLTSDCCDLASVLSASPSLKELDQONNLDVGRLLCGRLR-HPACKLIRLGLDQT 961
DB 793 KNRITSEGGKCVLAVNKTSTIVDVGWGNQIGDEGAKAFALAKDHPB--ITTLISLAFN 850
QY 962 TILSDENROEL-RALEQEKPLLIIFSRKPSVMTPTGLDTGEMSTSSLSKRLQRLGSE 1020
DB 851 GISPEGGKSLAQLAKNTTLTIV-----LTKNELNDESAECFAEMLRVNQT 897
QY 1021 ASHV 1024
DB 898 LRHL 901

RESULT 13
US-09-340-620A-43
; Sequence 43; Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US/09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US/09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US/09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US/09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-43

Query Match 4.6%; Score 348; DB 4; Length 953;
Best Local Similarity 21.6%; Pred. No. 3.4e-24;
Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

QY 234 WAAVGTTPQAHTSLQPHHPWPSVRESLCTWPKWNEFQKFTQLLLQRP-----287
DB 103 WLSEIGFSP-----SOLIRTKTIIVTDPVSRYTQQLRHQGRDSKFM 144
QY 288 --HPSQDPLVKRSPD-----YVENRGHLEIRDLF--GPGDLOEPRIVILQGA 336
DB 145 LCYAKEDLLLEETVMDTLMGLVGFNNENLGLDCLDHDSTGVINEHGETVFVFGDA 204
QY 337 GIGKSTLARQVKEAWRGQLYGDRFQHVYFSCRELA---QSKVSLAEELIGKD---GTA 390
DB 205 GVGKSMLLQRLQSLWASGRLTSTA-KFFHFRCRMFSCFKESDMLSLQDLLFKHFCYCEPQ 263
QY 391 TPAPTRQILSR-PERLLFILDGVD-----PGWVLOEPPSELCLHWSQPOPADALLGSLGK 446
DB 264 DPEEVFSFLRPHTALTFTDGLDELHSDFDLSRPVDS--CCPWEPAHPL-VLLANLISG 320
QY 447 TILPEASFLITARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYTDERQAIAPRL 506
DB 321 RLLKGAGKLLTARTGV---EVPRLRKKVLLRGFSPSHLRAYARMPPERTAQEHLLQ 377
QY 507 VSKNELMALCLVPWVSWLACTCLMQ-----SOLIRTKTIIVTDPVSRYTQQLRHQGRDSKFM 144
DB 378 LDANPNLCSLCGVPLFCWIIIFRCFOHQTVPFEGSSQLPDCAVTLTDVFLVTEVHLNRP 437
QY 538 EKLTLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWQKTLFSPDRLKRGHD 597
DB 438 QPSSLVQNRTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEVQASKLO 490
QY 598 GAIISFLKMGILQEHPP-----IPLSYSFHLCFOEFAAMSYLEDE-----640
DB 491 ---EGDLQGLFRALPDVGPQGGQYEFPHLTLOAFTAFPLVADDDKYSTRELLRFFRE 546
QY 641 ---KRGKHSNCIIDLKTLAYGHLGFGAS-----VPSL-----QQLLOPHSLESILHCLY 723
DB 547 WTSPEGATSSSCHSP-----FSPQCLGSRSLGPDPRNKHDFQFTNLFVCGLLAKAR 600
QY 683 EREMENIFHCRLSOGRLNMQW-----VPSL-----QQLLOPHSLESILHCLY 723
DB 601 QKLLRLQVPKAILRRKRKALWAHLFASLSRYLSKSLPRVQSGGFNQVHAMPTFLWMLRCIY 660
QY 724 ETRNKTFLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHKVKKLQLEGRQHRSTWSP 783

Db 661 ETQ-----SOKVGRLAARGI-----SADYKLAFCNACSDCSALSFLVHFRQL----- 706
QY 784 VVLFRRVPTDAYWQILFVLKTRNLKELDLSGNSLSHSAVKSLSCKTLRRPRCLLETIR 843
Db 707 -----ALDLDNNLNDYGVQELQPCFSR-----LTVIR 734
QY 844 LAGCGLTAECDKDLAFGLRANOTLTDLFNVLTADAGAKHLQORLPQSCCK-LORLQLV 902
Db 735 LSNVQITDGTGVKLCBELTKYKIVTFLGYNQITDIGARYVAQILDE--CRGLKHLKLG 792
QY 903 SCGLTSDCCDLASVLSASPSLKELDLOONLDDVGRVLLCEGLR-HPACKLIRLGLDOT 961
Db 793 KNRITSEGGCKVALAVKNSTSIVDVGMWGNQIGDEGAKAPAFALKDHPG--LTTLSLAFN 850
QY 962 TILSDMRQEL-RALEQEKPOLLIFFSRKPSVMTPTTEGLDTGMSNSTSLKQRQLGSERA 1020
Db 851 GISPEGKSLAQAOKNTTLTIW-----LYKNELDNDESAECFAEMLRVNOT 897
QY 1021 ASHV 1024
Db 898 LRHL 901

RESULT 14

US-08-519-547A-6

; Sequence 6, Application US/08519547A

; Patent No. 5994082

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Proteins Essential for the Expression of

; TITLE OF INVENTION: Vertebrate MHC Class II Genes, DNA Sequences Encoding Same

; TITLE OF INVENTION: and Pharmaceutical Compositions

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FISH & NEAVE

; STREET: 1251 AVENUE OF THE AMERICAS

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: U.S.A.

; ZIP: 10020-1104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/519,547A

; FILING DATE: 25-AUG-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP94113378.7

; FILING DATE: 26-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, JAMES F.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: VOS-11

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-596-9000

; TELEFAX: 212-596-9090

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1130 amino acids

; TYPE: amino acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-519-547A-6

Query Match

Best Local Similarity 4.6%; Score 345; DB 2; Length 1130;

Pred. No. 9.3e-24;

Matches 256; Conservative 134; Mismatches 430; Indels 282; Gaps 49;
QY 13 LEFLKE-----ELKEFOLLANKAHSRSSGETPA-----QPEKTSQMEVAS 55
Db 111 LEGLSKDIFKHIGDPDEVIGESMEMPAEVQSKQKPFPEELPADLKHWPABPPTVWTGS 170
QY 56 YLVAQYGORAWDLALHTWQMLRSLC-----AQAEAGAHSPSPYPSPSEPHLSPSP 111
Db 171 LLVGVSD-----CSTLPLPLPALFNQEPASQMRLEKTDQIPMPSSSLCLNLP 223
QY 112 TSTAVLMPWIELPAGCTGGERRV-----LRPLDPTSGRRMREISASLLY 157
Db 224 EGPIQFVTITSLPHGLWQISEAGTGVSSIFIYHGEVPOASQVPPPSG-----FTV 274
QY 158 QALPSPDHESPSQSNAPTSTAVLGSWGPQPSLAPR-----EQEAPGTQWLDSTG 213
Db 275 HGLPTSPDR--PGSTSPAPSATDL-----PSMPEPALTSRANWTEHKTSPTQCP--AAG 325
QY 214 IYYTEIREREREKSEKGRPPWAAVVGTPPQAITSLOPHHHPWEPSPVRESLCSCTWKNED 273
Db 326 -----EVSNK-LPKWPEPV---EQFYRSLQDTYGA-EPAGPDGI----- 359
QY 274 FNQKFTQLLLQRPHPRSQDPLVKE--SWPDVVEEN--RGHLIEIRDLFGPGLDQEPRI 329
Db 360 -----LVEVDLVQARLERSSSKSELERLATPDWAERQLAQGLAEVLLAAKEHRRPRETRV 415
QY 330 VILQGAAGIGKSTLARQVKEAWGRQLYGDRFQHYFYFSCRELAQ--SKVVSIAELIGKDG 388
Db 416 IAVLKAGQKSYWAGAVSRWACGRL--PQYDFVFSVPCNRPDGDYAGQDILLFSLG 473
QY 389 TATPAPI-----RQILSRPERLLFDGVDPGCVLQEPSSELCIHWSPQPAD----- 437
Db 474 ---PQPLVAADVEFVSHILKPRDRLVLLIDAPEE-----LEAQDGLHSTCGPAPAEPCSL 525
QY 438 -ALLCSLLGKTYLPEASFLITARTTALQNLISLSQA--RWVEVLGSESSRKEYFYRYFT 495
Db 526 RGLLAGLFQKLLRGCTLLLTARPRG--RLVQSLSKADALFELSGFSMEQAQYVNMRYFE 583
QY 496 D----ERQAIRAFRLVKGNKELWALCLVPWVSWLACTCLMQOMKREKLTLSKTTTTC 551
Db 584 SSGMTEHQD-RALTLLRDRPPLLSHSHSPTLCRAVCQSEALLELGEDAKLPS-TLTGLY 641
QY 552 LHYLAQALQAPGLQRLDCLSLAAE-GIWOKKTL-----FSPDDLKRGHGLDGAIISTFLK 606
Db 642 VGLLGRAALDSDPG-ALAEALAKLAWELGRHSHOSTLQEDQFPSADVR-----TWAMA 691
QY 607 MGILQHP-----IPLSY-SFHLCP-QEPPAAMSVLEDE-----KGRKHSNCI 650
Db 692 KGLVQHPRAESELAPFSFLQCFGLALWALSSEIKDKELPOYLALTPRKKRYPDNWL 751
QY 651 IDLEKTLBAYGIHGLFGASTTRFLLGLLSDEG-----EREM 686
Db 752 EGVPRFLA-----GLIFOPARCLGALLGPSAAASVDRKQKVLARYLKLQGTGLRQOL 806
QY 687 ENIFHC-----RLSQGRNLMQWVP-----SLQLLQPHSLR----- 717
Db 807 LELLHCAHEAREAGIWHVQVQLPGRSLFGLTRTPDAAHVLGKALEAAGQFSLDLRST 866
QY 718 -----SLHCLYTRNTFTTQWMAHFEEMCMVETDM-----ELIVCTFCIKF 760
Db 867 GICPSGLSLVGLSCV--TRFRAALSDTVLWESLRQHGETKLLQAAEKFTEIPEPKAS 924
QY 761 SRHVKKL-QLIEGRQHRSTWSPMTMVL-----FRWVPTDAY-WQILFSLVKVTR 808
Db 925 LKQVEDLGLKVQTRRSSSEDTAGELPAVRDLKLEFALGVPVSGPQAFPKVIRIITAFS 984
QY 809 NLKELD---LSCNSLSHSAVKSLSCKTLRRPRCLLETIRLAGCGLTAEDCKDLAFGLRANQ 865
Db 985 SLQHLDDLALSENKIGDEGVSQLSATF--PQL-----K 1015
QY 866 TTTELDLSPNVLTDAGAKHLQORLPQSCCKLQRLQVSCGLTSDCCQDLASVLSASPSLK 925
Db 1016 SLETLNLSQNNITDLGAYKLAELPSLAASLLRLSLYNNNCICDVGAESLARVLPDMVSLR 1075

QY 926 ELDLQNNLDDVGVRLCEGLR 947
 Db 1076 VMDVQYNKFTAGAAQQAASLR 1097
 RESULT 15
 US-09-207-359B-47
 ; Sequence 47, Application US/09207359B
 ; Patent No. 6469140
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE REFERENCE: 07334-112001
 ; CURRENT APPLICATION NUMBER: US/09/207,359B
 ; CURRENT FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 966
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: majority sequence
 US-09-207-359B-47

Query Match 4.3%; Score 326.5; DB 4; Length 966;
 Best Local Similarity 21.9%; Pred. No. 4.5e-22;
 Matches 187; Conservative 113; Mismatches 325; Indels 227; Gaps 31;
 QY 271 NEDFNQKFTQLLLQRP-----HPSQDPLVKSQWPD-----YVEENRGHLEIR 315
 Db 122 NTDPSRYTQQLRHQLGRDSKFVLCYAQKEDLLLEIYMDTLMLGLVGFNSIGSLGLA 181
 QY 316 DLFF--CPGLDTQEPRIVLQAGIGKSTLQVKEAWGRQL--VGDHFQHFVFSREL 372
 Db 182 CLLDHSTGVNLQGETVFLVGDAGVGSMLLQRLQSLWASGRUTAGAKF--FFHFRCRMF 239
 QY 373 A---QSKVSVLAELIGKD---GTATPAPIRQILSR--PERLLFILDGVDFPGWLOBPSS 425
 Db 240 SCFESDRLSLQDLLEKHFCEYEQDEEVFAFLRPHVAFVTFGLD-----ELHSD 292
 QY 426 LCLH-----WSQOPADALGSLGKLTILPEASFLITARTALQNLIPSLEQARWVE 477
 Db 293 LDLSRVDPDSSCPWEPAHPL--VLLANLLSGKLLKAGKLLTARTGV---EVPRQLLRKKVL 348
 QY 478 VLGPSESSRKEYFYRFTDERQAIRAFRLVKSNNELMALCLVPWVSWLACTCLMQ----- 532
 Db 349 LRGFSPSHLRVARRMFPERRAAQDHLLSOLDANPNLCISLGPVLCWIIIFRCFQHFQAAF 408
 QY 533 -----QMKRKEKLTLSKTTTLLC--LHYLAQALQAOPLGP 566
 Db 409 EGSSSQLPDCAVTLTDVFLVTEVHLNRMQPSLSVQNRTRSAPETLHAGRTIHA----- 463
 QY 567 QLRDLCSLAAEGTWQKTLFSPDDLRKGLDGAIIITFLKMGILQHP-----IPLSYS 620
 Db 464 ----LGEVAHRTGDKSLFVFGQEVQASGLQ---EGDLQLGFLRALPDVPGGQGSYE 515
 QY 621 FTHLCFQEEFAANSYVLEDEKGRGK-----HNCIIDLEKTLIAYG 661
 Db 516 FPHLTLOAPFTAFFLVADDKVGTELLRFFQEWTSFGGAASSCHSSFL-----SFQCLG 570
 QY 662 IHGLFGASTTRFLGLLSDEGEREMENIFHCLR---SQRNLMQWVPSL----- 707
 Db 571 GSGRAGED-----LFKNKHQFTNLFVCGLLAKAKQLLRQLVPAALRRKRKALWA 623
 QY 708 -----QLLOPHSLSLHCLYETRNKTFLTQYMAHFEEMGMCV 745

Search completed: January 29, 2004, 13:48:52
 Job time : 31.1576 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 ; Search time 249.019 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534
Sequence: 1 MAGGAWGRACYLEFLKKE.....HLIMELWEKSGKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7534	100.0	1429	10	US-09-996-617-2 Sequence 2, Appli
2	7534	100.0	1429	10	US-09-931-071-2 Sequence 2, Appli
3	7534	100.0	1429	12	US-10-028-374-15 Sequence 15, Appl
4	7534	100.0	1429	12	US-10-183-770-15 Sequence 15, Appl
5	7534	100.0	1429	15	US-10-028-392-11 Sequence 11, Appl
6	7502	99.6	1473	10	US-09-388-221-2 Sequence 2, Appli
7	7488	99.4	1429	12	US-10-028-374-3 Sequence 3, Appli
8	7488	99.4	1429	12	US-10-183-770-3 Sequence 3, Appli
9	7364	97.7	1399	10	US-09-388-221-4 Sequence 4, Appli
10	7332	97.3	1443	10	US-09-388-221-6 Sequence 6, Appli
11	6344	84.2	1454	10	US-09-388-221-10 Sequence 10, Appl
12	6174	81.9	1424	10	US-09-388-221-12 Sequence 12, Appl
13	4072	54.0	764	12	US-10-407-866-92 Sequence 92, Appl
14	2171.5	28.8	442	11	US-09-895-298-139 Sequence 139, App
15	1382.5	18.4	1027	12	US-10-407-866-68 Sequence 68, Appl

16	1380	18.3	1061	15	US-10-066-521-18	Sequence 18, Appl
17	1375	18.3	1035	11	US-09-965-621-24	Sequence 24, Appl
18	1375	18.3	1035	11	US-10-407-866-24	Sequence 24, Appl
19	1364	18.1	1004	12	US-10-108-260A-3161	Sequence 3161, Ap
20	1279	17.0	1034	12	US-10-028-374-18	Sequence 18, Appl
21	1279	17.0	1034	12	US-10-132-967-5	Sequence 5, Appli
22	1279	17.0	1034	12	US-10-183-770-18	Sequence 18, Appl
23	1279	17.0	1034	14	US-10-127-516-5	Sequence 5, Appli
24	1279	17.0	1034	14	US-10-027-629-5	Sequence 5, Appli
25	1154	15.3	896	15	US-10-066-521-22	Sequence 22, Appl
26	1037	13.8	994	11	US-09-965-621-16	Sequence 16, Appl
27	1037	13.8	994	12	US-10-407-866-16	Sequence 16, Appl
28	1037	13.8	994	15	US-10-066-521-24	Sequence 24, Appl
29	1008	13.4	919	12	US-10-094-749-2718	Sequence 2718, Ap
30	985.5	13.1	1162	15	US-10-216-645-2	Sequence 2, Appli
31	973.5	12.9	635	12	US-10-407-866-90	Sequence 90, Appl
32	972.5	12.9	858	9	US-09-848-035-8	Sequence 8, Appli
33	972.5	12.9	858	10	US-09-986-224-8	Sequence 4, Appli
34	969	12.9	1143	15	US-10-216-645-4	Sequence 4, Appli
35	948	12.6	1062	12	US-10-239-663-43	Sequence 43, Appl
36	945.5	12.5	1033	12	US-10-132-967-2	Sequence 2, Appli
37	945.5	12.5	1033	14	US-10-127-516-2	Sequence 2, Appli
38	945.5	12.5	1033	14	US-10-027-629-2	Sequence 2, Appli
39	945.5	12.5	1344	15	US-10-056-521-6	Sequence 6, Appli
40	943.5	12.5	1016	10	US-09-986-224-19	Sequence 19, Appl
41	941.5	12.5	1049	12	US-10-239-663-42	Sequence 42, Appl
42	924.5	12.3	732	12	US-10-407-866-72	Sequence 72, Appl
43	922.5	12.2	952	12	US-10-407-866-70	Sequence 70, Appl
44	913.5	12.1	674	12	US-10-407-866-89	Sequence 89, Appl
45	908.5	12.1	980	9	US-09-848-035-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-996-617-2
; Sequence 2, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-617-2

Query Match 100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGGAWGRACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ 60
DB	1	MAGGAWGRACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ 60
QY	61	YGEQAWDLALHTWQMGRLSLCAQAEAGAGHSFPYPSPSPHGLSPSQPTSTAVLMPW 120
DB	61	YGEQAWDLALHTWQMGRLSLCAQAEAGAGHSFPYPSPSPHGLSPSQPTSTAVLMPW 120
QY	121	IHELPAQCCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHSPSQSPNAPTST 180

Db 121 IHELPACTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQSNAPTST 180
Qy 181 AVLGWSGSPPOPSLAPREOEAPGTQWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Db 181 AVLGWSGSPPOPSLAPREOEAPGTQWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Qy 241 PPOAHTSLOPHHHHPWSPVRESLCSWPMKNEFNQKFTQLLLQORPHPRSDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPWSPVRESLCSWPMKNEFNQKFTQLLLQORPHPRSDPLVKRSW 300
Qy 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Qy 361 FOHVYFSCRELAQSKVSLAEILGDKGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Db 361 FOHVYFSCRELAQSKVSLAEILGDKGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Qy 421 EPSSELCHWSQOPADALLGSLGKTIILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Db 421 EPSSELCHWSQOPADALLGSLGKTIILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Qy 481 FSESSRKEYFYRFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Db 481 FSESSRKEYFYRFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQAOLPQLRDLSLAEGIWOKTFLFSPDRLKRGHLDGAI 600
Db 541 TLTSKTTTTLCHLYLAQAOLPQLRDLSLAEGIWOKTFLFSPDRLKRGHLDGAI 600
Qy 601 ISTFLKXGILQEHPIPLSYSFHLCQEPFAAMSVYLEDKGRGHSNCIIDLKTELEY 660
Db 601 ISTFLKXGILQEHPIPLSYSFHLCQEPFAAMSVYLEDKGRGHSNCIIDLKTELEY 660
Qy 661 GIHGLFGASTTFLGLLSDGEGRENIHFCHRLSQGRNLMQWVPSLQLLLOPHSLESJH 720
Db 661 GIHGLFGASTTFLGLLSDGEGRENIHFCHRLSQGRNLMQWVPSLQLLLOPHSLESJH 720
Qy 721 CLYETENKFTLVQMAHFEEMGCVETDMLLVCTEIKFHSRVKQLQIEGRQHRSTWS 780
Db 721 CLYETENKFTLVQMAHFEEMGCVETDMLLVCTEIKFHSRVKQLQIEGRQHRSTWS 780
Qy 781 PTMVVLFWRVVPVTDAYQWILFVSVLKYTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
Db 781 PTMVVLFWRVVPVTDAYQWILFVSVLKYTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
Qy 841 TURLACGLTAEDCKDLAFGLRANQTLTBLDLSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
Db 841 TURLACGLTAEDCKDLAFGLRANQTLTBLDLSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
Qy 901 LVSCGLTSCCQDLASVLSASPSLKBELDLQONNLDVGVRLICEGLRHHPACKLIRGLDQ 960
Db 901 LVSCGLTSCCQDLASVLSASPSLKBELDLQONNLDVGVRLICEGLRHHPACKLIRGLDQ 960
Qy 961 TTLSDSMRQELRALEOEKQOLLIFSRKPSVMTPTGLDTGEMSNSTSLKQRLGSERA 1020
Db 961 TTLSDSMRQELRALEOEKQOLLIFSRKPSVMTPTGLDTGEMSNSTSLKQRLGSERA 1020
Qy 1021 ASHVAQANLLDVSKEIPPIAETAEBSSEVVPVELLCVPSPASQGLDHTKPLGTDGDFW 1080
Db 1021 ASHVAQANLLDVSKEIPPIAETAEBSSEVVPVELLCVPSPASQGLDHTKPLGTDGDFW 1080
Qy 1081 GPTGPVATEVDEKELNLYRHVPVAGSYRWPNTGLCFVNRVATVIEIFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVDEKELNLYRHVPVAGSYRWPNTGLCFVNRVATVIEIFCVWDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLDIIKAEPAVAEAVHLPHFVALQGGHVDTSLFQMAHFEKEGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLDIIKAEPAVAEAVHLPHFVALQGGHVDTSLFQMAHFEKEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSPSLGVLLKMHNAALRFIPVTSVLLYHVRHPEEVTTHLYLIPSDCSIR 1260

Db 1201 ELHHIVLENPSPSLGVLLKMHNAALRFIPVTSVLLYHVRHPEEVTTHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGBDQLFSEFYVGHLSGIRLVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGBDQLFSEFYVGHLSGIRLVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLGQVLSQEQYVERVLAENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLDKLGQVLSQEQYVERVLAENTRPS 1380
Qy 1381 QMRKLFSLQSQWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLQSQWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429

RESULT 2
US-09-931-071-2
; Sequence 2, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PXT
; ORGANISM: Homo sapiens
US-09-931-071-2

Query Match 100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGGAWGRACYLEPLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSQMEVASYLVAQ 60
Db 1 MAGGAWGRACYLEPLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSQMEVASYLVAQ 60
Qy 61 YGEQAWDLAHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
Db 61 YGEQAWDLAHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
Qy 121 THELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQSNAPTST 180
Db 121 THELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSSPDHESPSQSNAPTST 180
Qy 181 AVLGWSGSPPOPSLAPREOEAPGTQWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Db 181 AVLGWSGSPPOPSLAPREOEAPGTQWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Qy 241 PPOAHTSLOPHHHHPWSPVRESLCSWPMKNEFNQKFTQLLLQORPHPRSDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPWSPVRESLCSWPMKNEFNQKFTQLLLQORPHPRSDPLVKRSW 300
Qy 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Qy 361 FOHVYFSCRELAQSKVSLAEILGDKGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Db 361 FOHVYFSCRELAQSKVSLAEILGDKGTATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Qy 421 EPSSELCHWSQOPADALLGSLGKTIILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Db 421 EPSSELCHWSQOPADALLGSLGKTIILPEASFLITARTTALQNLIPSLQARWVEVLG 480

QY	481	FSESSRKEYFYRYFTDERQAI	RAFR	LVKSNKELWALCLVPWVSWLACTCLMQMKEKEL	540
DB	481	FSESSRKEYFYRYFTDERQAI	RAFR	LVKSNKELWALCLVPWVSWLACTCLMQMKEKEL	540
QY	541	TLTSTKTTTTLCHYLAAQALQ	APLGPQ	LDLCSLAAEGIWOKTFLSPDDLRKHGLDGA	600
DB	541	TLTSTKTTTTLCHYLAAQALQ	APLGPQ	LDLCSLAAEGIWOKTFLSPDDLRKHGLDGA	600
QY	601	ISTFLKMGILQEHPIPLSYF	HLCPQEF	FAAMSVYVLEDEKGRGHSNCIIDLKLEAY	660
DB	601	ISTFLKMGILQEHPIPLSYF	HLCPQEF	FAAMSVYVLEDEKGRGHSNCIIDLKLEAY	660
QY	661	GIHGLFGASTTRFLGLLSD	DEGEREMENIFH	CHLSQGRNLMQWVPSLQLLQPHSLES	720
DB	661	GIHGLFGASTTRFLGLLSD	DEGEREMENIFH	CHLSQGRNLMQWVPSLQLLQPHSLES	720
QY	721	CLYETRNKFTLTQVMAHFE	MGCMVETD	MELLVCTCFIKFSRHVKQLIEGRQHR	780
DB	721	CLYETRNKFTLTQVMAHFE	MGCMVETD	MELLVCTCFIKFSRHVKQLIEGRQHR	780
QY	781	PTMVLFWRVPVTDAYWQIL	FVSVLKV	TRNLKELDLSGNSLSHSAVSKLCTLR	840
DB	781	PTMVLFWRVPVTDAYWQIL	FVSVLKV	TRNLKELDLSGNSLSHSAVSKLCTLR	840
QY	841	TLRACGLTAEDCKDLAFGL	RANQTLT	ELDLSFNVLTDAGAKHLQRLQPSCKL	900
DB	841	TLRACGLTAEDCKDLAFGL	RANQTLT	ELDLSFNVLTDAGAKHLQRLQPSCKL	900
QY	901	LVSCGLTSDCCDLASVLS	ASPSL	KELDLQNNLDVGVRLLCGHRHACKL	960
DB	901	LVSCGLTSDCCDLASVLS	ASPSL	KELDLQNNLDVGVRLLCGHRHACKL	960
QY	961	TTLSDEMRELALQEKQOL	IFSRKPS	VMTPTTEGLDTGENSNTSSIKRQL	1020
DB	961	TTLSDEMRELALQEKQOL	IFSRKPS	VMTPTTEGLDTGENSNTSSIKRQL	1020
QY	1021	ASHVAQANLKLLDVSKI	FPFAE	IAESSPEVVPVELLCVSPASQGD	1080
DB	1021	ASHVAQANLKLLDVSKI	FPFAE	IAESSPEVVPVELLCVSPASQGD	1080
QY	1081	GPTGPVATEVVDKKNLYR	VHFFVAG	SYRPNWGLCFVMEAVTVIEECVMD	1140
DB	1081	GPTGPVATEVVDKKNLYR	VHFFVAG	SYRPNWGLCFVMEAVTVIEECVMD	1140
QY	1141	POHSMVAGPLLDIKAE	PGVAE	AVHLPHFVALGCHVDTSLFQMAH	1200
DB	1141	POHSMVAGPLLDIKAE	PGVAE	AVHLPHFVALGCHVDTSLFQMAH	1200
QY	1201	ELHHIVLENPSFPLG	VLLKMIH	NALRFIPVTSVLLYHRVHPEEVT	1260
DB	1201	ELHHIVLENPSFPLG	VLLKMIH	NALRFIPVTSVLLYHRVHPEEVT	1260
QY	1261	KELECYRSPGEDQL	SEFVYV	GHLSGIRLOVKDKDELTVWALYK	1320
DB	1261	KELECYRSPGEDQL	SEFVYV	GHLSGIRLOVKDKDELTVWALYK	1320
QY	1321	ARTAVSPIDAPOLLH	VDQYR	EQRLARTVSVEVLDKLHGQVLSQ	1380
DB	1321	ARTAVSPIDAPOLLH	VDQYR	EQRLARTVSVEVLDKLHGQVLSQ	1380
QY	1381	QMRKFLSLSQSWDR	CKKQGLYQ	ALKETHPHLIMELMEKSGKGLPLSS	1429
DB	1381	QMRKFLSLSQSWDR	CKKQGLYQ	ALKETHPHLIMELMEKSGKGLPLSS	1429

RESULT 3
US-10-028-374-15
; Sequence 15, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRBM1

```

; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-15

Query Match      100.0%; Score 7534; DB 12; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWGRACLVLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSMEVASLYVAQ 60
DB 1 MAGGAWGRACLVLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSMEVASLYVAQ 60

QY 61 YGEQAWDLALHTWQMGRLSLCAQAEQAGHSPSPFPYSPSEPHLGSPOPTSTAVLMPW 120
DB 61 YGEQAWDLALHTWQMGRLSLCAQAEQAGHSPSPFPYSPSEPHLGSPOPTSTAVLMPW 120

QY 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDPHESPOESPNAPTST 180
DB 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDPHESPOESPNAPTST 180

QY 181 AVLGSWGSPQPSLAPREQEAFTGWPDLDETSGIYYTEIREREKSEKGRPPWAAVVG 240
DB 181 AVLGSWGSPQPSLAPREQEAFTGWPDLDETSGIYYTEIREREKSEKGRPPWAAVVG 240

QY 241 PPQATSLQPHHHPWEPSPVRESLCSTWPKNEDFNOKTQLLLQLRPHSPSDPLVKRSW 300
DB 241 PPQATSLQPHHHPWEPSPVRESLCSTWPKNEDFNOKTQLLLQLRPHSPSDPLVKRSW 300

QY 301 PDYVEENRHGLIEIRDLFGPLDTPQPRIVILQGAAGICKSTLAROVKEAWGQOLYGD 360
DB 301 PDYVEENRHGLIEIRDLFGPLDTPQPRIVILQGAAGICKSTLAROVKEAWGQOLYGD 360

QY 361 FOHVYFSCRELAQSKVSLAELIGDKGTATAPRIOLSRPERLLFILDGVDEPGWVLQ 420
DB 361 FOHVYFSCRELAQSKVSLAELIGDKGTATAPRIOLSRPERLLFILDGVDEPGWVLQ 420

QY 421 EPSSSELCHWSQOPADALLSGLSLGTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
DB 421 EPSSSELCHWSQOPADALLSGLSLGTILPEASFLITARTTALQNLIPSLQARWVEVLG 480

QY 481 FSESSRKEYFYRYFTDERQAI RAFRLVKNKELWALCLVPWVSWLACTCLMQMKEKEL 540
DB 481 FSESSRKEYFYRYFTDERQAI RAFRLVKNKELWALCLVPWVSWLACTCLMQMKEKEL 540

QY 541 TLTSKTTTTLCHYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDL RKHGLDGA 600
DB 541 TLTSKTTTTLCHYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDL RKHGLDGA 600

QY 601 ISTFLKMGILQEHPIPLSYFHLCPQEFPAAMSVYVLEDEKGRGHSNCIIDLKLEAY 660
DB 601 ISTFLKMGILQEHPIPLSYFHLCPQEFPAAMSVYVLEDEKGRGHSNCIIDLKLEAY 660

QY 661 GIHGLFGASTTRFLGLLSDDEGEREMENIFHCHLSQGRNLMQWVPSLQLLQPHSLES 720
DB 661 GIHGLFGASTTRFLGLLSDDEGEREMENIFHCHLSQGRNLMQWVPSLQLLQPHSLES 720

QY 721 CLYETRNKFTLTQVMAHFEEMGCMVETD MELLVCTCFIKFSRHVKQLIEGRQHRSTWS 780
DB 721 CLYETRNKFTLTQVMAHFEEMGCMVETD MELLVCTCFIKFSRHVKQLIEGRQHRSTWS 780

QY 781 PTMVLFWRVPVTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
DB 781 PTMVLFWRVPVTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840

```

FILE REFERENCE: D0067NP
CURRENT APPLICATION NUMBER: US/10/028,374
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/257,773
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 1429
TYPE: PRT
ORGANISM: homo sapiens
US-10-028-374-15

Query Match 100.0%; Score 7534; DB 12; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLSDSNVLTDAAGAKHLQRLQPSCKLQRLQ 900
841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLSDSNVLTDAAGAKHLQRLQPSCKLQRLQ 900
901 LVSCGLTSDCCQDLASVLASPSLKELDLQNNLDDVGVRLLCEGLRHHPACKLIRLGLDQ 960
901 LVSCGLTSDCCQDLASVLASPSLKELDLQNNLDDVGVRLLCEGLRHHPACKLIRLGLDQ 960
961 TTLSDEMROELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSESA 1020
961 TTLSDEMROELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSESA 1020
1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEPVVVELLCVPSASQDHLTKPLGDDDFW 1080
1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEPVVVELLCVPSASQDHLTKPLGDDDFW 1080
1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVWDQFLGEIN 1140
1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVWDQFLGEIN 1140
1141 POHSHMMVAGPLDDIKAEPCGAVAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
1141 POHSHMMVAGPLDDIKAEPCGAVAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
1201 ELHHIVLENPSPGLVLLKMIHNLRFIPVTSVLLYHRVHPEEVTFLYLIIPSDCSIR 1260
1201 ELHHIVLENPSPGLVLLKMIHNLRFIPVTSVLLYHRVHPEEVTFLYLIIPSDCSIR 1260
1261 KELELCYRSPGSDQFSEFVYVGHLSGIRLQVKDKDETFLWEALVKGDLMPATTLIPP 1320
1261 KELELCYRSPGSDQFSEFVYVGHLSGIRLQVKDKDETFLWEALVKGDLMPATTLIPP 1320
1321 ARIAVPSPLDAPQLLHFVDVQYREOLIAVTSVEVVDLKLHGQVLSQEOYERVLAENTRPS 1380
1321 ARIAVPSPLDAPQLLHFVDVQYREOLIAVTSVEVVDLKLHGQVLSQEOYERVLAENTRPS 1380
1381 QMRKLFSLSQSDWRCKCKGQLYQALKETHPHLIMELWEKSKKGLLPLSS 1429
1381 QMRKLFSLSQSDWRCKCKGQLYQALKETHPHLIMELWEKSKKGLLPLSS 1429

RESULT 4
US-10-183-770-15
; Sequence 15, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRB1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-15

Query Match 100.0%; Score 7534; DB 12; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGGAWGRACLYLEFLKKEELKEFQALLANKAHSRSSGSETPAQPEKTSQMEVASYLVAQ 60
1 MAGGAWGRACLYLEFLKKEELKEFQALLANKAHSRSSGSETPAQPEKTSQMEVASYLVAQ 60
61 YGEQRAWDLALHTWEQMLRSILCAQAEAGAGHSPPYPSPSEPHLGSPQSTSTAVLMPW 1200

61 YGEQRAWDLALHTWEQMLRSILCAQAEAGAGHSPPYPSPSEPHLGSPQSTSTAVLMPW 120
121 THEIPAGCTCGSERVLPQDPTSGRRWRETSASLLYQALPSSPDHSPSPESNPATST 180
121 THEIPAGCTCGSERVLPQDPTSGRRWRETSASLLYQALPSSPDHSPSPESNPATST 180
181 AVLGSWGSPPPSLAPRQEQEAPGTQWPLDETSGIYYTIREREREKSKGRPPMAAVVGT 240
181 AVLGSWGSPPPSLAPRQEQEAPGTQWPLDETSGIYYTIREREREKSKGRPPMAAVVGT 240
241 PPQATSLQPHHPWPSPVRESLSTWPKNEDFNQKFTQLLLQRPHPRSQDPLVKRSW 300
241 PPQATSLQPHHPWPSPVRESLSTWPKNEDFNQKFTQLLLQRPHPRSQDPLVKRSW 300
301 PDYVEENRGHLIEIRDLFGPGLDQEPRIIVILQGAAGIGKSTLAKQVKEAMGRGOLYDGR 360
301 PDYVEENRGHLIEIRDLFGPGLDQEPRIIVILQGAAGIGKSTLAKQVKEAMGRGOLYDGR 360
361 FQHVYFSCRELAOSKVVSIAELIGKDGATAPAPIQILSRPERLLFLIDGVDEFGWLQ 420
361 FQHVYFSCRELAOSKVVSIAELIGKDGATAPAPIQILSRPERLLFLIDGVDEFGWLQ 420
421 EPSSCLLHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSLQARWVEVLG 480
421 EPSSCLLHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSLQARWVEVLG 480
481 FSESRKRYFYRYFTDERQAIAPRLVKSNEKELWALCLVPMVSWLACTCLMQOMKREKEL 540
481 FSESRKRYFYRYFTDERQAIAPRLVKSNEKELWALCLVPMVSWLACTCLMQOMKREKEL 540
541 TLTSKTTTTLCHVLAQALQAPLQPOLRDICSLAAEGIWQKTLFSPDDLKHLGDGAI 600
541 TLTSKTTTTLCHVLAQALQAPLQPOLRDICSLAAEGIWQKTLFSPDDLKHLGDGAI 600
601 ISTFLKMGILQEHPIPLSYFHLICQFQFFAAMSYYLDEKGRGHKSNCCIIDLEKTLBAY 660
601 ISTFLKMGILQEHPIPLSYFHLICQFQFFAAMSYYLDEKGRGHKSNCCIIDLEKTLBAY 660
661 GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSOGNLMQWVPSLQLLQPHSLESILH 720
661 GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSOGNLMQWVPSLQLLQPHSLESILH 720
721 CLYETRNKTFITQVMAHPEEMGMVETDMLLVCTFCIKFSRHVKKQLIIEGRQHRSTWS 780
721 CLYETRNKTFITQVMAHPEEMGMVETDMLLVCTFCIKFSRHVKKQLIIEGRQHRSTWS 780
781 PTMVVLPFRWVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE 840
781 PTMVVLPFRWVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE 840
841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLSDSNVLTDAAGAKHLQRLQPSCKLQRLQ 900
841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLSDSNVLTDAAGAKHLQRLQPSCKLQRLQ 900
901 LVSCGLTSDCCQDLASVLASPSLKELDLQNNLDDVGVRLLCEGLRHHPACKLIRLGLDQ 960
901 LVSCGLTSDCCQDLASVLASPSLKELDLQNNLDDVGVRLLCEGLRHHPACKLIRLGLDQ 960
961 TTLSDEMROELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSESA 1020
961 TTLSDEMROELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSESA 1020
1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEPVVVELLCVPSASQDHLTKPLGDDDFW 1080
1021 ASHVAQANLKLDDVSKIFPIAETAEBSSEPVVVELLCVPSASQDHLTKPLGDDDFW 1080
1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVWDQFLGEIN 1140
1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVNRREAVTVIEFCVWDQFLGEIN 1140
1141 POHSHMMVAGPLDDIKAEPCGAVAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
1141 POHSHMMVAGPLDDIKAEPCGAVAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200

Db 1141 POHSMWVAGPLLDIAKAPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLSEKPARV 1200
Qy 1201 ELHHIVLENPSPFGLVLLKMIHNAIRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPFGLVLLKMIHNAIRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGSDQDLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPTATLIPP 1320
Db 1261 KELELCYRSPGSDQDLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPTATLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLHGQVLSQEOYERVLAEINTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLHGQVLSQEOYERVLAEINTRPS 1380
Qy 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429

RESULT 5
US-10-028-392-11
; Sequence 11, Application US/10028392
; Publication No. US20030087340A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN NERVOUS SYSTEM TISSUES, HLRRNS1
; FILE REFERENCE: D0085.02
; CURRENT APPLICATION NUMBER: US/10/028,392
; PRIORITY FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/259,479
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US 60/260,616
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 11
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-392-11

Query Match 100.0%; Score 7534; DB 15; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGGAWGRACLYLEFLKXELKEFOLLANKAHSRSSGETPAQPKTSGMEVASLYVAQ 60
Db 1 MAGGAWGRACLYLEFLKXELKEFOLLANKAHSRSSGETPAQPKTSGMEVASLYVAQ 60
Qy 61 YGEQRAWDALHTWQWGLRSLCAQAEAGAGHSPSPYSPSEPHLGSPOPTSTAVLMPW 120
Db 61 YGEQRAWDALHTWQWGLRSLCAQAEAGAGHSPSPYSPSEPHLGSPOPTSTAVLMPW 120
Qy 121 IHLEPAGCTQGSERRVRLQPLDTSGRWRREISASLLYQALPSSPDHSPSPNAPTST 180
Db 121 IHLEPAGCTQGSERRVRLQPLDTSGRWRREISASLLYQALPSSPDHSPSPNAPTST 180
Qy 181 AVLGWSGPPQPSLAPREAPQTPWLDTSYIYTEIREREREKSEKGRPPWAAVGT 240
Db 181 AVLGWSGPPQPSLAPREAPQTPWLDTSYIYTEIREREREKSEKGRPPWAAVGT 240
Qy 241 PPOAHTSLOPHHPWPSPRESLCSWPMKNEFQKFTOLLORPHRSQDPLVKRSM 300
Db 241 PPOAHTSLOPHHPWPSPRESLCSWPMKNEFQKFTOLLORPHRSQDPLVKRSM 300
Qy 301 PDVVEENRHLIIRDLFGPLDTQEPRIVLQGAAGIGKSTLQVKEAWGRGQLYGR 360
Db 301 PDVVEENRHLIIRDLFGPLDTQEPRIVLQGAAGIGKSTLQVKEAWGRGQLYGR 360
Qy 361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDPEGWLIQ 420
Db 361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDPEGWLIQ 420

RESULT 6

US-09-388-221-2

; Sequence 2, Application US/09388221A

Qy 421 EPSSELCLHWSQPADALLGSLGKTTILPEASFLITARTTALQNLIPSLSEQARVVEVLG 480
Db 421 EPSSELCLHWSQPADALLGSLGKTTILPEASFLITARTTALQNLIPSLSEQARVVEVLG 480
Qy 481 PSESRKEFYRYFYFDEROARAPRLVKSNEKELWALCLVPVWSWLACTCLMOQMKRKEKL 540
Db 481 PSESRKEFYRYFYFDEROARAPRLVKSNEKELWALCLVPVWSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCLHYLAQALQAOPLQPOLRDLCSLAAEGIWOKKTLFSPDDLRKHGIDGAI 600
Db 541 TLTSKTTTTLCLHYLAQALQAOPLQPOLRDLCSLAAEGIWOKKTLFSPDDLRKHGIDGAI 600
Qy 601 ISTFLKMGILQEHPIPLSYFHLQCFQFFAAMS VVLEDEKGRGHKSNCCIIDEKTLLEY 660
Db 601 ISTFLKMGILQEHPIPLSYFHLQCFQFFAAMS VVLEDEKGRGHKSNCCIIDEKTLLEY 660
Qy 661 GIHGLFGASTTRFLGILGILSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHLSLH 720
Db 661 GIHGLFGASTTRFLGILGILSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHLSLH 720
Qy 721 CLYETRNKTFITQVMAHPEEMGCMVETDMLLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780
Db 721 CLYETRNKTFITQVMAHPEEMGCMVETDMLLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780
Qy 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
Db 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
Qy 841 TURLAGCGLTABDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
Db 841 TURLAGCGLTABDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
Qy 901 LVSCGLTSDCCODLASVLSASPSLAKELDLQNNLDDVGVRLCEGLRHHPACKLIRLGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPSLAKELDLQNNLDDVGVRLCEGLRHHPACKLIRLGLDQ 960
Qy 961 TTLSDEMQRQELRAEQEKPOLLI FSRKRPVSMVTPTTEGLDTGEMSNSTSLKQRQLSERA 1020
Db 961 TTLSDEMQRQELRAEQEKPOLLI FSRKRPVSMVTPTTEGLDTGEMSNSTSLKQRQLSERA 1020
Qy 1021 ASHVAQANLKLDDVSKIPIPIAIESSPEVVPVELLCVPSPASQDGLHTKPLGTTDDDFW 1080
Db 1021 ASHVAQANLKLDDVSKIPIPIAIESSPEVVPVELLCVPSPASQDGLHTKPLGTTDDDFW 1080
Qy 1081 GPTGPVATEVVDKKNLYRVHPFPVAGSYRWPNMTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKKNLYRVHPFPVAGSYRWPNMTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLLDIAKAPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLSEKPARV 1200
Db 1141 PQHSMWVAGPLLDIAKAPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLSEKPARV 1200
Qy 1201 ELHHIVLENPSPFGLVLLKMIHNAIRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPFGLVLLKMIHNAIRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGSDQDLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPTATLIPP 1320
Db 1261 KELELCYRSPGSDQDLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPTATLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLHGQVLSQEOYERVLAEINTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTVTSVEVLDKLHGQVLSQEOYERVLAEINTRPS 1380
Qy 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429

Publication No. US20020192643A1
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-2

Query Match 99.6%; Score 7502; DB 10; Length 1473;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAGGAWRLACYLEFLKKEBELKEFOLLANKAHSRSSGETPAQPEKTSGMVEASVLAQ 60
DB 1 MAGGAWRLACYLEFLKKEBELKEFOLLANKAHSRSSGETPAQPEKTSGMVEASVLAQ 60
QY 61 YGEQRAWDALHTWEQMLRSICAQAEAGAGHSPSPFYPSPHLSGSPQPTSTAVLMPW 120
DB 61 YGEQRAWDALHTWEQMLRSICAQAEAGAGHSPSPFYPSPHLSGSPQPTSTAVLMPW 120
QY 121 IHELPAGTCQSGRRVRLQIPDTSGRWRREISASLLYQALPSPDHESQSPNAPTST 180
DB 121 IHELPAGTCQSGRRVRLQIPDTSGRWRREISASLLYQALPSPDHESQSPNAPTST 180
QY 181 AVLGSGSPPOPSLAPREGEAPCTQPLDTSIGYITEIREREREKSEKGRPPWAAVGT 240
DB 181 AVLGSGSPPOPSLAPREGEAPCTQPLDTSIGYITEIREREREKSEKGRPPWAAVGT 240
QY 241 PPOAHTSLOPHHPWPSPVRESLCSWPWKNEFDNQFTQLLLQRPSPQDPLVKRSW 300
DB 241 PPOAHTSLOPHHPWPSPVRESLCSWPWKNEFDNQFTQLLLQRPSPQDPLVKRSW 300
QY 301 PDYVEENRGLHIBIRLFGGLDTPRIVLQGAAGIKSTLARQVKEAWGEGQLYGR 360
DB 301 PDYVEENRGLHIBIRLFGGLDTPRIVLQGAAGIKSTLARQVKEAWGEGQLYGR 360
QY 361 FOHFVFSRELAQSKVSLAEILGKGTATPAIRQILSRPERLLFILDGVDPEQWVLQ 420
DB 361 FOHFVFSRELAQSKVSLAEILGKGTATPAIRQILSRPERLLFILDGVDPEQWVLQ 420
QY 421 EPSSELCHWSQPQADALLGSLGKTLIPASFLITARTTALQNLIPSLQARWVEVLG 480
DB 421 EPSSELCHWSQPQADALLGSLGKTLIPASFLITARTTALQNLIPSLQARWVEVLG 480
QY 481 FSESSRKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
DB 481 FSESSRKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
QY 541 TLTSKTTTTLCHLYLAQAQPLGQPLDRDLCSLAAGIWKQKTLFSPDRLKHGLDGA 600
DB 541 TLTSKTTTTLCHLYLAQAQPLGQPLDRDLCSLAAGIWKQKTLFSPDRLKHGLDGA 600
QY 601 ISTFLXWGLQEHPIPLSYSFHLCQFEFFAAMSYYVLEDEKGRGHSNCIIDLKTELEY 660
DB 601 ISTFLXWGLQEHPIPLSYSFHLCQFEFFAAMSYYVLEDEKGRGHSNCIIDLKTELEY 660
QY 661 GIHGLFGASTTRPLGLLSDGGERENI FHCRLSQGRNLMQWVPSQLLLQPHSLESJH 720
DB 661 GIHGLFGASTTRPLGLLSDGGERENI FHCRLSQGRNLMQWVPSQLLLQPHSLESJH 720
QY 721 CLYETRNKFTLVQMAHFEEMGCMVETDMELVCTCFIKFSRHVKQLQIEGRQHSSTWS 780
DB 721 CLYETRNKFTLVQMAHFEEMGCMVETDMELVCTCFIKFSRHVKQLQIEGRQHSSTWS 780
QY 781 PTMVVLFWRVPVTDAYWQILFVSLKTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840

DB 781 PTMVVLFWRVPVTDAYWQILFVSLKTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
QY 841 TLRLAGCGLTABDCXDLAFGLRANOTLTJELDLSFNVLTDAGAKHLCQRLRQPSCKLRLQ 900
DB 841 TLRLAGCGLTABDCXDLAFGLRANOTLTJELDLSFNVLTDAGAKHLCQRLRQPSCKLRLQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKDLDQNNLDVGVRLCEGLRHLPACKLIRGLDQ 960
DB 901 LVSCGLTSDCCODLASVLSASPSLKDLDQNNLDVGVRLCEGLRHLPACKLIRGLDQ 960
QY 961 TTLSDEMRQELRALRQEKYQPOLLI FSRKRPSTVMTPTTEGLDTGEMSNSTSLKQRUGSRA 1020
DB 961 TTLSDEMRQELRALRQEKYQPOLLI FSRKRPSTVMTPTTEGLDTGEMSNSTSLKQRUGSRA 1020
QY 1021 ASHVAQANLKLLDVSKIPIPIAEIAESSPEVVPVVELLCVPSASQGLHTRKPLGTDDDFW 1080
DB 1021 ASHVAQANLKLLDVSKIPIPIAEIAESSPEVVPVVELLCVPSASQGLHTRKPLGTDDDFW 1080
QY 1081 GFTGPVATEVVDKEKNLYRVHFPVAGSYRWPNGLCFVMREAVTVIEIFCVWDQFLGEIN 1140
DB 1081 GFTGPVATEVVDKEKNLYRVHFPVAGSYRWPNGLCFVMREAVTVIEIFCVWDQFLGEIN 1140
QY 1141 PQSHMWAGPLLDIKABFGAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPARV 1200
DB 1141 PQSHMWAGPLLDIKABFGAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPARV 1200
QY 1201 EUHHIVLENPSFPLGVLLKMHNLARFIPVTSVLLVHRVHPBEVTHLYLIPSDCSIR 1260
DB 1201 EUHHIVLENPSFPLGVLLKMHNLARFIPVTSVLLVHRVHPBEVTHLYLIPSDCSIR 1260
QY 1261 -----KELELCYRSPEGDQLF 1276
DB 1261 KAIDDLKMKFQVRTHKPPPLTPLYMGCRYTVSGSGSGMLLEILPKLELCYRSPEGDQLF 1320
QY 1277 SFYVGHGSGIRLQVKDKDETLWBEALVKPGDLMPTATTLIPPARIAVPSPLDAPOLLH 1336
DB 1321 SFYVGHGSGIRLQVKDKDETLWBEALVKPGDLMPTATTLIPPARIAVPSPLDAPOLLH 1380
QY 1337 FVDQYREQLIARVTSVEVVDLKHQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1396
DB 1381 FVDQYREQLIARVTSVEVVDLKHQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1440
QY 1397 KDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
DB 1441 KDGLYQALKETHPHILMELWEKSGKGLPLSS 1473

RESULT 7
US-10-028-374-3
; Sequence 3, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-3

Query Match 99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWRLACYLEFLKKEBELKEFOLLANKAHSRSSGETPAQPEKTSGMVEASVLAQ 60

Db 1 MAGGAWRLACYLEFLKKEELKEFQALLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ 60
QY 61 YGEQRAWDALHTWEQMLRSLSLCAQAEAGAGHSPFPYSPSEPHLGSPQPTSTAVLMPW 120
Db 61 YGEQRAWDALHTWEQMLRSLSLCAQAEAGAGHSPFPYSPSEPHLGSPQPTSTAVLMPW 120
QY 121 IHELPAQCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHSPQSPNAPTST 180
Db 121 IHELPAQCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHSPQSPNAPTST 180
QY 181 AVLGSNGSPPOPSLAPREOEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
Db 181 AVLGSNGSPPOPSLAPREOEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
QY 241 PPOAHTSLOPHHHHPSPVRESLCSWPMKNEFQKFTQLLLQRPSPQDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPSPVRESLCSWPMKNEFQKFTQLLLQRPSPQDPLVKRSW 300
QY 301 PDYVEENRGLHLEIRDLFGGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHLEIRDLFGGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
QY 361 FQHVYFSCRELAQSKVSLAEILGKGTATPAPIRQILSRPERILFILDGVDPEQWVLQ 420
Db 361 FQHVYFSCRELAQSKVSLAEILGKGTATPAPIRQILSRPERILFILDGVDPEQWVLQ 420
QY 421 EPSSELCLHWSQOPADALLGSLGKTLTILPEASFLITARTALONLIPSLQARWVEVLG 480
Db 421 EPSSELCLHWSQOPADALLGSLGKTLTILPEASFLITARTALONLIPSLQARWVEVLG 480
QY 481 FSESSRKEYFYRYFTDERQAIRFLVKNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
Db 481 FSESSRKEYFYRYFTDERQAIRFLVKNKELWALCLVPWVSWLACTCLMQQMKRKEKL 540
QY 541 TLTSKTTTILCLHYLAQAQPLGQPLRDLCSLAABEGIWQKTLFSPDDLRKHGLDGA 600
Db 541 TLTSKTTTILCLHYLAQAQPLGQPLRDLCSLAABEGIWQKTLFSPDDLRKHGLDGA 600
QY 601 ISTFLKMGILQEHPIPLVSFIHLFCQEPFAAMSVYLEDKGRGKHSNCIIDLKLEAY 660
Db 601 ISTFLKMGILQEHPIPLVSFIHLFCQEPFAAMSVYLEDKGRGKHSNCIIDLKLEAY 660
QY 661 GIHLFGASTTRFLGLLSDEGEREMENIFCHRLSQGRNLMQWVPSLQLLQPHSLSLH 720
Db 661 GIHLFGASTTRFLGLLSDEGEREMENIFCHRLSQGRNLMQWVPSLQLLQPHSLSLH 720
QY 721 CLYETRNKPTLTQMAHFEEMGMCVETDMELLVCTCFIKPSRHVKKQLQIEGRQHRSTWS 780
Db 721 CLYETRNKPTLTQMAHFEEMGMCVETDMELLVCTCFIKPSRHVKKQLQIEGRQHRSTWS 780
QY 781 PTMVVLFPMVPTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
Db 781 PSMVVLFRMVPTDAYWQILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
QY 841 TRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
Db 841 TRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLMDAGAKHLQORLQPSCKLQRLQ 900
QY 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDGVVRLLCCEGLRHHPACKILRGLDQ 960
Db 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDGVVRLLCCEGLRHHPACKILRGLDQ 960
QY 961 TTLSDEMQRLEALEQKPOLLI FSRKPSVMTPTGLDGENSNTSSSLKROLGSE 1020
Db 961 TTLSDEMQRLEALEQKPOLLI FSRKPSVMTPTGLDGENSNTSSSLKROLGSE 1020
QY 1021 ASHVAQANLKLDVSKIFPTAEIAESSPEWVVELLCVPSASQDGLHKTPTGTDDEFW 1080
Db 1021 ASHVAQANLKLDVSKIFPTAEIAESSPEWVVELLCVPSASQDGLHKTPTGTDDEFW 1080
QY 1081 GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVMEAVTVIEFCVWDQFLGEIN 1140

Db 1081 GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVMEAVTVIEFCVWDQFLGEIN 1140
QY 1141 POHSWMVAGPLDLDIKAEPCGAVEAVHLPHFEVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Db 1141 POHSWMVAGPLDLDIKAEPCGAVEAVHLPHFEVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
QY 1201 ELHHIVLENPSPSLGVLKMIHNAIRFTPTVTSVLLYHRVHPEEVTFFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPSLGVLKMIHNAIRFTPTVTSVLLYHRVHPEEVTFFHLYLIPSDCSIR 1260
QY 1261 KELELCYRSPGDDOLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGDDOLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
QY 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTSVSEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARTSVSEVVLDKLHGQVLSQEQYERVLAENTRPS 1380
QY 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSGKGLPLSS 1429

RESULT 8
US-10-183-770-3
; Sequence 3, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: PREDOMINATELY IN BONE MARROW, HLRREMI
; CURRENT APPLICATION NUMBER: US/10/183,770
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-3

Query Match 99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAGGAWRLACYLEFLKKEELKEFQALLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ 60
Db 1 MAGGAWRLACYLEFLKKEELKEFQALLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ 60
QY 61 YGEQRAWDALHTWEQMLRSLSLCAQAEAGAGHSPFPYSPSEPHLGSPQPTSTAVLMPW 120
Db 61 YGEQRAWDALHTWEQMLRSLSLCAQAEAGAGHSPFPYSPSEPHLGSPQPTSTAVLMPW 120
QY 121 IHELPAQCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHSPQSPNAPTST 180
Db 121 IHELPAQCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHSPQSPNAPTST 180
QY 181 AVLGSNGSPPOPSLAPREOEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
Db 181 AVLGSNGSPPOPSLAPREOEAPCTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
QY 241 PPOAHTSLOPHHHHPSPVRESLCSWPMKNEFQKFTQLLLQRPSPQDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPSPVRESLCSWPMKNEFQKFTQLLLQRPSPQDPLVKRSW 300
QY 301 PDYVEENRGLHLEIRDLFGGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHLEIRDLFGGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGDR 360

361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Db
361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Qy 421 EPSSELCHWSQPPADALLGSLGKTLITPEASFLITARTTALQNLIPSELAQWVEVLG 480
Db 421 EPSSELCHWSQPPADALLGSLGKTLITPEASFLITARTTALQNLIPSELAQWVEVLG 480
Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSYFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLKTL 660
Db 601 ISTFLKMGILQEHPIPLSYFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLKTL 660
Qy 661 GHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
Db 661 GHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
Qy 721 CLYETRNTKFTLQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Db 721 CLYETRNTKFTLQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Qy 781 PTMVLFRWVPVTDAYQWILFSLVYKTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
Db 781 PSMVLFRWVPVTDAYQWILFSLVYKTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
Qy 841 TLRLACGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
Db 841 TLRLACGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLQORLQPSCKLQRLQ 900
Qy 901 LVSCGTLSDCCDLASVLSASPSLKELDIQNNLDVGVRLLCGLRHHPACKLIRGLDQ 960
Db 901 LVSCGTLSDCCDLASVLSASPSLKELDIQNNLDVGVRLLCGLRHHPACKLIRGLDQ 960
Qy 961 TTLSDEMRLQALAEQEKQLLIFSRKPSVMTPTTEGLDTGEMSNSTSLKRLQSGERA 1020
Db 961 TTLSDEMRLQALAEQEKQLLIFSRKPSVMTPTTEGLDTGEMSNSTSLKRLQSGERA 1020
Qy 1021 ASHVAQANLKLVDKIPPIAEIAESSPEVVPVELLCVPSASQGDHLTKPLGTTDDPW 1080
Db 1021 ASHVAQANLKLVDKIPPIAEIAESSPEVVPVELLCVPSASQGDHLTKPLGTTDDPW 1080
Qy 1081 GPTGPVATEVVDKKNLYRVHPPVAGSYRWPNTGLCFVMEAVTVBEFCVMDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKKNLYRVHPPVAGSYRWPNTGLCFVMEAVTVBEFCVMDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLFOMAFKEEGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLFOMAFKEEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSPLGVLKMIHNAIRFTPVTSVLLYHRVHPEEVTFLHYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPLGVLKMIHNAIRFTPVTSVLLYHRVHPEEVTFLHYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGEDOLFSEFVGHGSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGEDOLFSEFVGHGSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Qy 1321 ARIAPVSPDLAPOLLHFVQYREQLIARTVTSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
Db 1321 ARIAPVSPDLAPOLLHFVQYREQLIARTVTSVEVLDKLHGQVLSQEQYERVLAENTRPS 1380
Qy 1381 QMRKLFSLSQSWDRKCKDGLYQALKEKTHPHLMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRKCKDGLYQALKEKTHPHLMELWEKSGKGLPLSS 1429

RESULT 9
US-09-388-221-4
; Sequence 4, Application US/09388221A
; Publication NO. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-4

Query Match 97.7%; Score 7364; DB 10; Length 1399;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 30; Gaps 1;
Qy 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
Db 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSMEVASYLVAQ 60
Qy 61 YGEQAWDLALHTWQMGILRSILCAQAEAGAGHSPSPFPYSPSEPHLGSPQSTSTAVLMPW 120
Db 61 YGEQAWDLALHTWQMGILRSILCAQAEAGAGHSPSPFPYSPSEPHLGSPQSTSTAVLMPW 120
Qy 121 HELPAGCTQGSERVLRLQPLDTSRRWRREISASLLYQALPSSPDHSPSQESNPATST 180
Db 121 HELPAGCTQGSERVLRLQPLDTSRRWRREISASLLYQALPSSPDHSPSQESNPATST 180
Qy 181 AVLSGWSGPPQPSLAPREQEAPGTQPLDETSGIYVYTERERERERERERERERERERER 240
Db 181 AVLSGWSGPPQPSLAPREQEAPGTQPLDETSGIYVYTERERERERERERERERERERER 240
Qy 241 PQAHSTLQPHHHPEVPSVRESLCSSTWPKNEDFNOKTQOLLQRPSPQDPLVKRWSW 300
Db 241 PQAHSTLQPHHHPEVPSVRESLCSSTWPKNEDFNOKTQOLLQRPSPQDPLVKRWSW 300
Qy 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAMGRGQLYGR 360
Db 301 PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAMGRGQLYGR 360
Qy 361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Db 361 FOHVYFSCRELAQSKVSLAEIIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ 420
Qy 421 EPSSELCHWSQPPADALLGSLGKTLITPEASFLITARTTALQNLIPSELAQWVEVLG 480
Db 421 EPSSELCHWSQPPADALLGSLGKTLITPEASFLITARTTALQNLIPSELAQWVEVLG 480
Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQALQAPLGPQLRDLCSLAAEGIWOKTFLSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSYFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLKTL 660
Db 601 ISTFLKMGILQEHPIPLSYFIHLCFQEFPAAMSYVLEDEKGRGKHSNCIIDLKTL 660
Qy 661 GHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
Db 661 GHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES 720
Qy 721 CLYETRNTKFTLQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780

Db	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS	780
Qy	781	PTMVLFWRVPVTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Db	781	PTMVLFWRVPVTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Qy	841	TURLAGCGLTAECDKDLAFGLRANQTLFELDLFNVLTADAGAKHLCORLRQPSCKLRLQ	900
Db	841	TURLAGCGLTAECDKDLAFGLRANQTLFELDLFNVLTADAGAKHLCORLRQPSCKLRLQ	900
Qy	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHHPACKLIRLGLDQ	960
Db	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHHPACKLIRLGLDQ	960
Qy	961	TTLSDMRQELRALRQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRQLGSE	1020
Db	958	-----KPSVMTPTTEGLDTGEMSNSTSLKQRQLGSE	990
Qy	1021	ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW	1080
Db	991	ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW	1050
Qy	1081	GPTGPVATEVDKKNLVRVHPFVAGSVRWPNTGLCFVNRVAVTVEIRFCVWDQFLGEIN	1140
Db	1051	GPTGPVATEVDKKNLVRVHPFVAGSVRWPNTGLCFVNRVAVTVEIRFCVWDQFLGEIN	1110
Qy	1141	POHSMWVAGPLLDIKAEFGAVEAVHLPHFVALQGGHVDTSLSFQMAHFEKGMLEKPARV	1200
Db	1111	POHSMWVAGPLLDIKAEFGAVEAVHLPHFVALQGGHVDTSLSFQMAHFEKGMLEKPARV	1170
Qy	1201	ELHHIVLENPSPLGVLLKMHNLALRFTPVTSVLLYHRVHPEEVTTHLYLIPSDCSIR	1260
Db	1171	ELHHIVLENPSPLGVLLKMHNLALRFTPVTSVLLYHRVHPEEVTTHLYLIPSDCSIR	1230
Qy	1261	KELELCYRSPGDELQFSEFVYVGHGSGIRLOVKDKDETLVWEALVKPGDLMPATLIPP	1320
Db	1231	KELELCYRSPGDELQFSEFVYVGHGSGIRLOVKDKDETLVWEALVKPGDLMPATLIPP	1290
Qy	1321	ARIAVPSPLDAPQLLHFVDQYREQIARTVTSVEVLDKLGQVLSQEQYERVLAEINTRPS	1380
Db	1291	ARIAVPSPLDAPQLLHFVDQYREQIARTVTSVEVLDKLGQVLSQEQYERVLAEINTRPS	1350
Qy	1381	QMRKLFSLQSQWDRCKCGLYQALKEHPLHIMELWEKSKGLPLSS	1429
Db	1351	QMRKLFSLQSQWDRCKCGLYQALKEHPLHIMELWEKSKGLPLSS	1399

RESULT 10

US-09-388-221-6
; Sequence 6, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-6

Query Match 97.3%; Score 7332; DB 10; Length 1443;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 74; Gaps 2;
Qy 1 MAGGAWGLRACYLEFLKKEBELKEFOLLANKAHSRSSSGTTPAQPKTSGMEVASVLAQ 60
Db 1 MAGGAWGLRACYLEFLKKEBELKEFOLLANKAHSRSSSGTTPAQPKTSGMEVASVLAQ 60

Qy	61	YGEQAWDLALHTWEOMGLRSICAQOEGAGHSPPYSPSEPHLGSPSQPTSTAVLMPW	120
Db	61	YGEQAWDLALHTWEOMGLRSICAQOEGAGHSPPYSPSEPHLGSPSQPTSTAVLMPW	120
Qy	121	IHELPAAGTCQSERVRLQPLDTSGRWRREISASLLYQALPSSPDHESPSQESPNAPTST	180
Db	121	IHELPAAGTCQSERVRLQPLDTSGRWRREISASLLYQALPSSPDHESPSQESPNAPTST	180
Qy	181	AVLSGWSGPPQPSLAPREQEAPGTOWPLDETSGIYYTIRERERERESKGRPPMAVVT	240
Db	181	AVLSGWSGPPQPSLAPREQEAPGTOWPLDETSGIYYTIRERERERESKGRPPMAVVT	240
Qy	241	PPQHTSLQPHHPWPSPVRESLCSLTPWKNEDFNKFTQLLLQRPSPRSDPLVKRSM	300
Db	241	PPQHTSLQPHHPWPSPVRESLCSLTPWKNEDFNKFTQLLLQRPSPRSDPLVKRSM	300
Qy	301	PDYVENRGLHLEIRDLFGPLDTPRIIVILQGAAGIKSTLARQVKEAWGRGQLYGR	360
Db	301	PDYVENRGLHLEIRDLFGPLDTPRIIVILQGAAGIKSTLARQVKEAWGRGQLYGR	360
Qy	361	FQHVFFYFSCRELAQSKVSVLAELIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ	420
Db	361	FQHVFFYFSCRELAQSKVSVLAELIGKDGATATPAPIRQILSRPERLLFILDGVDEPGWVLQ	420
Qy	421	EPSSCLLHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSLEQARVVEVLG	480
Db	421	EPSSCLLHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSLEQARVVEVLG	480
Qy	481	PSESRKKEVYFYFTDERQATRAPRLVKSNNELWALCLVPVWSWLACTCLMOMKKEKL	540
Db	481	PSESRKKEVYFYFTDERQATRAPRLVKSNNELWALCLVPVWSWLACTCLMOMKKEKL	540
Qy	541	TILSTTTTTLCLHYLAQALQAPLQPLQDLCSLAAEGIWQKKTILFSPDDLKKGIDGAI	600
Db	541	TILSTTTTTLCLHYLAQALQAPLQPLQDLCSLAAEGIWQKKTILFSPDDLKKGIDGAI	600
Qy	601	ISTFLKMGILQHPITPLSYFIHLQFQFFAAMSIVLEDEKGRGHNSCIIIDLEKTLAY	660
Db	601	ISTFLKMGILQHPITPLSYFIHLQFQFFAAMSIVLEDEKGRGHNSCIIIDLEKTLAY	660
Qy	661	GIHGLFGASTTRFLGLLSDDEGEREMENIFHCLRSQGRNLMQWPSLQLLLOPHLSLH	720
Db	661	GIHGLFGASTTRFLGLLSDDEGEREMENIFHCLRSQGRNLMQWPSLQLLLOPHLSLH	720
Qy	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS	780
Db	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS	780
Qy	781	PTMVLFWRVPVTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Db	781	PTMVLFWRVPVTDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Qy	841	TURLAGCGLTAECDKDLAFGLRANQTLFELDLFNVLTADAGAKHLCORLRQPSCKLRLQ	900
Db	841	TURLAGCGLTAECDKDLAFGLRANQTLFELDLFNVLTADAGAKHLCORLRQPSCKLRLQ	900
Qy	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHHPACKLIRLGLDQ	960
Db	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHHPACKLIRLGLDQ	960
Qy	961	TTLSDMRQELRALRQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRQLGSE	1020
Db	958	-----KPSVMTPTTEGLDTGEMSNSTSLKQRQLGSE	990
Qy	1021	ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW	1080
Db	991	ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW	1050
Qy	1081	GPTGPVATEVDKKNLVRVHPFVAGSVRWPNTGLCFVNRVAVTVEIRFCVWDQFLGEIN	1140
Db	1051	GPTGPVATEVDKKNLVRVHPFVAGSVRWPNTGLCFVNRVAVTVEIRFCVWDQFLGEIN	1110
Qy	1141	POHSMWVAGPLLDIKAEFGAVEAVHLPHFVALQGGHVDTSLSFQMAHFEKGMLEKPARV	1200

Db 1430 ALDVLFRSISERDPYLVSYL 1449

RESULT 12

US-09-388-221-12

Sequence 12, Application US/09388221A

Publication No. US20020192643A1

GENERAL INFORMATION:

APPLICANT: Reed, John C.

TITLE OF INVENTION: PAD Domain-Containing Polypeptides,

FILE REFERENCE: P-LJ 3650

CURRENT APPLICATION NUMBER: US/09/388, 221A

CURRENT FILING DATE: 1999-09-01

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 12

LENGTH: 1424

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-388-221-12

Query Match 81.9%; Score 6174; DB 10; Length 1424;

Best Local Similarity 82.1%; Pred. No. 0;

Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

QY 1 MAGGAWGRACYLEFLFKKEELKEFOLLANKAHSRSSSGETPAQPEKTSGMVEASVYLAQ 60

Db 1 MAGGAWGRACYLEFLFKKEELKEFOLLANKAHSRSSSGETPAQPEKTSGMVEASVYLAQ 60

QY 61 YGEQAWDLALHTWEQMLRSICAQAEAGAGHSPFPYPSPHLSGSPQPTSTAVLMPW 120

Db 61 YGEQAWDLALHTWEQMLRSICAQAEAGAGHSPFPYPSPHLSGSPQPTSTAVLMPW 120

QY 121 IHLPAGCTQGSERRVRLDPTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180

Db 121 IHLPAGCTQGSERRVRLDPTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180

QY 181 AVLGSWGPQPSLAPREOAGCTOWPLDETSGIYYTEIREREREKSGRPPMAAVGT 240

Db 181 AVLGSWGPQPSLAPREOAGCTOWPLDETSGIYYTEIREREREKSGRPPMAAVGT 240

QY 241 PQAHTSLQPHHPWEPSPRESLCTWPKNEDFNQKFTQLLLQRPSPRSDPLVKRSW 300

Db 241 PQAHTSLQPHHPWEPSPRESLCTWPKNEDFNQKFTQLLLQRPSPRSDPLVKRSW 300

QY 301 PUYVENRGLHTEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360

Db 301 PUYVENRGLHTEIRDLFGPGLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360

QY 361 FOHVYFSCRELAQSKVSLAEIIGKDGATAPRQILSRPERLLFILDGVDEPGWLQ 420

Db 361 FOHVYFSCRELAQSKVSLAEIIGKDGATAPRQILSRPERLLFILDGVDEPGWLQ 420

QY 421 EPSSELCLHWSQPQADALLGSLGKLTILPEASFLITARTTALQNLIPSEARWVEVLG 480

Db 421 EPSSELCLHWSQPQADALLGSLGKLTILPEASFLITARTTALQNLIPSEARWVEVLG 480

QY 481 FSESSRKEYFYRYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540

Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540

QY 541 TLTSKTTTTLCHLYLAQALQAPGLQDLRLCSLAAEGIWOKTLPSPDRLKHGLDGA 600

Db 541 TLTSKTTTTLCHLYLAQALQAPGLQDLRLCSLAAEGIWOKTLPSPDRLKHGLDGA 600

QY 601 ISTFLKMGILQEHPIPLSYFTHLCQEFFAAMSVYLEDEKGRGHSNCIIIDLEKTLAY 660

Db 601 ISTFLKMGILQEHPIPLSYFTHLCQEFFAAMSVYLEDEKGRGHSNCIIIDLEKTLAY 660

QY 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESJH 720

Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESJH 720

QY 721 CLYETRNTKTLTQYMAHFEENGCMVETDMLLVCTFCIKFSRHKVKKLQLEBGRQHRSTWS 780

Db 721 CLYETRNTKTLTQYMAHFEENGCMVETDMLLVCTFCIKFSRHKVKKLQLEBGRQHRSTWS 780

QY 781 PTMVVLFWRVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLSCKTLRRPRCLLE 840

Db 781 PTMVVLFWRVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLSCKTLRRPRCLLE 840

QY 841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLQPSCKLQRLQ 900

Db 841 TLRLAGCGLTAECDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLCORLQPSCKLQRLQ 900

QY 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960

Db 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960

QY 961 TTLSDEMQRLEALAEQEPQLLIFSRKRPSPVMTPEGLDGTGEMSNSTSSLRKQRLGSE 1020

Db 961 TTLSDEMQRLEALAEQEPQLLIFSRKRPSPVMTPEGLDGTGEMSNSTSSLRKQRLGSE 1020

QY 991 ASHVAQANLKLDDVSKIPPIAEIAEESSEPVVVELLCVPSASQGDHHTKPLGTDGDFW 1080

Db 991 ASHVAQANLKLDDVSKIPPIAEIAEESSEPVVVELLCVPSASQGDHHTKPLGTDGDFW 1080

QY 1081 GPTGVATVVDVKEKNLYRVHPFVAGSYRWENTGLCFVMREAVTVEIEBFCVMDQFLG-EI 1139

Db 1081 GPTGVATVVDVKEKNLYRVHPFVAGSYRWENTGLCFVMREAVTVEIEBFCVMDQFLG-EI 1139

QY 1140 NPQSHMMVAGPLDIDKAEF-GAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPA 1198

Db 1140 NPQSHMMVAGPLDIDKAEF-GAVEAVHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPA 1198

QY 1199 RVELHHVLENPSPGLVLLKMHNLARFIPVTSVLLYHRVHPEEVTFLYLIPSDCS 1258

Db 1199 RVELHHVLENPSPGLVLLKMHNLARFIPVTSVLLYHRVHPEEVTFLYLIPSDCS 1258

QY 1259 IR-----KELELCYRSPGEDQLF 1276

Db 1259 IR-----KELELCYRSPGEDQLF 1276

QY 1277 SEFYVGHGSGIRLOVKDKDETLLVMEALVKPGDLMPTTILIPARIAPVPSPLDAPQLLH 1336

Db 1277 SEFYVGHGSGIRLOVKDKDETLLVMEALVKPGDLMPTTILIPARIAPVPSPLDAPQLLH 1336

QY 1337 FVDQYREQLIARVTSVEVVLKHLH-GQVLSOEQYERVAENRTPSQMRKLSLSQSWDRK 1395

Db 1337 FVDQYREQLIARVTSVEVVLKHLH-GQVLSOEQYERVAENRTPSQMRKLSLSQSWDRK 1395

QY 1396 CKDGLYQALKEHPLHIMEL 1415

Db 1396 CKDGLYQALKEHPLHIMEL 1415

QY 1400 ALDVLFRSISERDPYLVSYL 1419

Db 1400 ALDVLFRSISERDPYLVSYL 1419

RESULT 13

US-10-407-866-92

Sequence 92, Application US/10407866

Publication No. US20040002593A1

GENERAL INFORMATION:

APPLICANT: Reed, John C.

TITLE OF INVENTION: PAD Domain-Containing Polypeptides,

FILE REFERENCE: 66654-10 (LJ 5755)

CURRENT APPLICATION NUMBER: US/10/407, 866

CURRENT FILING DATE: 2003-04-04

PRIOR APPLICATION NUMBER: US 60/370,538

PRIOR FILING DATE: 2002-04-04

NUMBER OF SEQ ID NOS: 129

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 92

LENGTH: 764

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-92

Query Match
Best Local Similarity 54.0%; Score 4072; DB 12; Length 764;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGAGWRLACYLEFLKKEBELKEFOLLANKAHSRSSGETPAQPKTSGMEVASYLVAQ 60
DB 1 MAGAGWRLACYLEFLKKEBELKEFOLLANKAHSRSSGETPAQPKTSGMEVASYLVAQ 60

QY 61 YGEQAWDLALHTWEQMLSLCAQAEAGAGHSPPYPSPHGLSPSQPTSTAVLMPW 120
DB 61 YGEQAWDLALHTWEQMLSLCAQAEAGAGHSPPYPSPHGLSPSQPTSTAVLMPW 120

QY 121 IHELPAGCTQGERRVRLQPLDTSGRWREISASLLYQALPSPDHESQSPNAPTST 180
DB 121 IHELPAGCTQGERRVRLQPLDTSGRWREISASLLYQALPSPDHESQSPNAPTST 180

QY 181 AVLGWGSPPQSLAPREQAPCTQWPLDETSGIYVTEIREREREKSGRPPMAAVGCT 240
DB 181 AVLGWGSPPQSLAPREQAPCTQWPLDETSGIYVTEIREREREKSGRPPMAAVGCT 240

QY 241 PQAHTSLOPHHHWPSPVRESLCSWPWKNEDEFNQKFTQLLLQRPHPRSQDPLVKRW 300
DB 241 PQAHTSLOPHHHWPSPVRESLCSWPWKNEDEFNQKFTQLLLQRPHPRSQDPLVKRW 300

QY 301 PDYVENRGLHIEIRDLFGGLDQTPRIVILQGAAGIKSTLARQVKAWGRGQLYGR 360
DB 301 PDYVENRGLHIEIRDLFGGLDQTPRIVILQGAAGIKSTLARQVKAWGRGQLYGR 360

QY 361 FOHVYFSCRELAQSKVSLAEIIGKDGATAPIRQILSRPERLLFIIDGVDPEGWLQ 420
DB 361 FOHVYFSCRELAQSKVSLAEIIGKDGATAPIRQILSRPERLLFIIDGVDPEGWLQ 420

QY 421 EPSSELCHWSQPQADALLGSLGKTLILPEASFLITARTTALQNLIPSEQARWVEVLG 480
DB 421 EPSSELCHWSQPQADALLGSLGKTLILPEASFLITARTTALQNLIPSEQARWVEVLG 480

QY 481 FSESSKEYFYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
DB 481 FSESSKEYFYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540

QY 541 TLTSKTTTTLCHLYLAQAQPLQPLQDLCLSLAAEGIWOKTFLSPDRLKHGLDGA 600
DB 541 TLTSKTTTTLCHLYLAQAQPLQPLQDLCLSLAAEGIWOKTFLSPDRLKHGLDGA 600

QY 601 ISTFLKMGILQEHPIPLSYSFHLCQEPFAAMSVYVLEDEKGRGHSNCIIDLKTL 660
DB 601 ISTFLKMGILQEHPIPLSYSFHLCQEPFAAMSVYVLEDEKGRGHSNCIIDLKTL 660

QY 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSL 720
DB 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSL 720

QY 721 CLYETENKTLQVMAHFEEMGVCVETDMLLVCTFCIKFSRHV 764
DB 721 CLYETENKTLQVMAHFEEMGVCVETDMLLVCTFCIKFSRHV 764

RESULT 14
US-09-895-298-139
; Sequence 139, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-68

Query Match
Best Local Similarity 28.8%; Score 2171.5; DB 11; Length 442;
Matches 414; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 997 GLDTGEMSNSTSSLKQRLGSEAAASHVAQANLKLDDVSKIFPIAETAEESPEV 1056
DB 1 GLDTGEMSNSTSSLKQRLGSEAAASHVAQANLKLDDVSKIFPIAETAEESPEV 60

QY 1057 LCVPSPASQGLDHTKPLGTDDDFWGPCTGPVATEVVDKKNLYRVHFPVAGSY 1116
DB 61 LCMPSPASQGLDHTKPLGTDDDFWGPCTGPVATEVVDKKNLYRVHFPVAGSY 120

QY 1117 FVMREAVTVEIEFCVWDQFLGEINPQHSWMVAGPLDIDKABPGAVEAVHLPH 1176
DB 121 FVMREAVTVEIEFCVWDQFLGEINPQHSWMVAGPLDIDKABPGAVEAVHLPH 180

QY 1177 VDTSLFQAHFKEEGMLLEKPARVELHHIVLENFSFSLGVLKMIHNAIRFIPV 1236
DB 181 VDTSLFQAHFKEEGMLLEKPARVELHHIVLENFSFSLGVLKMIHNAIRFIPV 240

QY 1237 LYHRVHPEVTHLYLIPSDCSIRKELELCYRSPGEDOLFSEFYVGHGSGIRLO 1296
DB 241 LYHRVHPEVTHLYLIPSDCSIRKELELCYRSPGEDOLFSEFYVGHGSGIRLO 300

QY 1297 DETLWAEALVPGDLMPATTILPPARIAPSPDAPQLLHFVDQYREQLIARVTS 1356
DB 301 DETLWAEALVPGDLMPATTILPPARIAPSPDAPQLLHFVDQYREQLIARVTS 360

QY 1357 DKLHGQVLSQBYQYERVAENRTPSQMRKLFSLQSDWCKDKGLYQALKETHPH 1416
DB 361 DKLHGQVLSQBYQYERVAENRTPSQMRKLFSLQSDWCKDKGLYQALKETHPH 417

QY 1417 EKGSK 1421
DB 418 NSGR 422

RESULT 15
US-10-407-866-68
; Sequence 68, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: 66654-10 (1J 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 1027
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-68

```

Search completed: January 29, 2004, 13:54:37
Job time : 252.019 secs

QY	8	RLACYLEFLKKEBELKEPOLLANKAHSRSSGTPP-AOPEKTSMEVASVLYAQYGEORA	66
Db	12	RLSTYLEEAEVLEAKFKPLYGTA--TELGEKIPMGSMKAGPLEMAQLLIHTHFGPEEA	69
QY	67	WDLALHTWEOQMGLRSLCAOQEGAGHSPSPYSPSEPHLGSPSO-PTSTAVLMPWIHELP	125
Db	70	WRLALSTFERINRKDLWERCQREDLVYDTPFGGPSS--LGNQSTCLLEVSIVTP--RKDP	125
QY	126	AGTCQSERVRLVQLPDTSGRRWREISLASLYQALPSSPDHESPOSQBNAPTSTAVLGS	185
Db	126	QETYROYVRKFKFLMEDRNARLCECVNLSHRYTRLLLVKEHSNPMQVQ-----	173
QY	186	WGSPPQSLAPRQEAPGTOWPLDETSGIYYTEIREREREKSEKGRPPWAAVVGTPPOAH	245
Db	174	-----	173
QY	246	TSLOPHHHHPWSPVRESLCSWPWKNEBQKFTQLLLQRPFRSQDPLVKRSPWDYVE	305
Db	174	-----QQLLDTGRGHARTV-----	187
QY	306	ENRGHL---IEIRDLPGLDQOE-PRIVILOGAAGIKSTLARQVKEANGRGOLYDRF	361
Db	188	---GHOASPIKIEITLFEPEDEPERPETTVMOGAAGIKSMIAKHVNLMDWADGKLFQGRF	244
QY	362	QHVYFSCRELAOSKV-VSLAEILGKDGATPAPIROIILSRPERLLFILQGVDEPGWVLQ	420
Db	245	DYLFYINCREMOSATECSMODLIFSCWPPSPAPLOELIRVPERLLFIIOGFBELKPSFH	304
QY	421	EPSSELCHWSQPQADALLGSLGKTLTPEASFLITARTALONLIPSLQEARWVEVLG	480
Db	305	DPOGPMCLWEEKRPTLELNSLIRKKLLPELSLLITRPTALEKHLRLEHPRHVEILG	364
QY	481	FSSSSKEYEYRYFTDEROAIRAFRLVKSNKELWALCLVPWVSWLACTCLMOQMKREKL	540
Db	365	FSAERKEYEYKYFHNAEQGVFNYYRDNPEUFTMCFVLVCMVVCTCLOQLEGGGLL	424
QY	541	TLTSKTTTTLCLHYLAQALQAPLGPOL-----RDLCSLAAGIMOKTKTLFSPDDLKRX	594
Db	425	RQTSRTTAVYMLYLLSLMQPKGAPRLQPPNQRGRLCSLAADGLMNOKILFEBQDLKRX	484
QY	595	GLDCAIISTPLKMGILQEH-PILPUSYFIHLCPQEPFAAMSYYL-EDEKGRGHNSCIID	652
Db	485	GLDGEDVSAPFLNNIIFOKINCERYYFIHLSFOEPPFAAMYIILDEGEGAGPDQ-----	540
QY	653	LEKTLBAYGI-HGLFGASTTRFLLGLLSDGEREMENIFHCRLSQ--GRNLMOWPSLQ-	708
Db	541	VTRLLEYASERSFLATSRFLFGLNETHRSHLEKSLCWKYSPHIKMDLLQWISKAQ	600
QY	709	---LLLOPHSLESILCHLYETRNKTLTQVMAHFEEMGC-VETDMELLVCTFCIKFSRHV	764
Db	601	SDGSTLQQSLEFPFSCLYIEQEBEFIQOALSHFOVIVSNIAKMEHWVSFCLKRCRSA	660
QY	765	KKLQLI-----EGQHRSTWSPTMVULFR--WVPVTAQWILFESVLKVYENLXELD	814
Db	661	QVILHYLGATYSAGEDRARCSAGAHULLVQLRPERTVILLDAYSBLHAAALCTNPMLIELS	720
QY	815	LSGNSLSHSAVKSICKTLRRPCILLETRLLAGCGLTAECDKDLAFGLRANOTLTDELDSF	874
Db	721	LYRNALGSRGVKLLCOGLRHPNCKLNLKRCRISSSACEDLSAALIANKNLTRMDLSG	780
QY	875	NVLTDAKAKHLQORLROPSCKLQRLQVLVSCGLTSDCCODLJASVLASPSLUKELDLQNNL	934
Db	781	NGVGFPGMMLLCEGLRHPQCRLOWIQLRKQLESGAQEWASVLGTNPHVLVELDLTGNAL	840
QY	935	DDVGVRLLCGLRHPACKLIRLGI-----DQTLSDMEQERALEQEKPOLLIIF	984
Db	841	EDJGLRLLCQGLRHPVCRKTLMLUKILRLTAACDELAELASTLSVNQSLRELDLSNEL---	897
QY	985	SRKPSVMTTEGLDTCGEMSNSTSSLRQRJGSEFAASHV-----AQANIKLLDVS	1035

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 2662.52 Seconds
(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccgctgcagcgggg.....atacgaagccagcttgaa 740

Scoring table: IDENTITY_NUC
Gapop 10_0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*			
1:	gb_ba.*	740	6 AR256267
2:	gb_htg.*	740	6 AX082246
3:	gb_in.*	740	6 AX536208
4:	gb_om.*	740	6 AX536223
5:	gb_ov.*	740	6 AF384665
6:	gb_pat.*	740	6 AX118619
7:	gb_ph.*	740	6 AF184073
8:	gb_pl.*	740	6 AX459863
9:	gb_pr.*	740	6 AB023416
10:	gb_ro.*	740	6 AX017270
11:	gb_scs.*	740	6 AX524974
12:	gb_sy.*	740	6 BD134441
13:	gb_un.*	740	6 BD057255
14:	gb_vl.*	740	6 AK000211
15:	em_ba.*	712	9 AF310103
16:	em_fun.*	712	9 BC013569
17:	em_hum.*	712	9 BC013569
18:	em_in.*	712	9 BC013569
19:	em_mu.*	712	9 BC013569
20:	em_om.*	712	9 BC013569
21:	em_or.*	712	9 BC013569
22:	em_ov.*	712	9 BC013569
23:	em_pat.*	712	9 BC013569
24:	em_ph.*	712	9 BC013569
25:	em_pl.*	712	9 BC013569
26:	em_ro.*	712	9 BC013569
27:	em_scs.*	712	9 BC013569
28:	em_un.*	712	9 BC013569
29:	em_vl.*	712	9 BC013569
30:	em_htg_hum.*	712	9 BC013569
31:	em_htg_inv.*	712	9 BC013569
32:	em_htg_other.*	712	9 BC013569
33:	em_htg_mus.*	712	9 BC013569
34:	em_htg_pln.*	712	9 BC013569
35:	em_htg_rod.*	712	9 BC013569
36:	em_htg_mam.*	712	9 BC013569
37:	em_htg_vrt.*	712	9 BC013569
38:	em_sy.*	712	9 BC013569
39:	em_htgo_hum.*	712	9 BC013569
40:	em_htgo_mus.*	712	9 BC013569
41:	em_htgo_other.*	712	9 BC013569

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	740	100.0	740	6	AR256267	AR256267 Sequence
2	740	100.0	740	6	AX082246	AX082246 Sequence
3	740	100.0	740	6	AX536208	AX536208 Sequence
4	740	100.0	740	6	AX536223	AX536223 Sequence
5	740	100.0	740	6	AF384665	AF384665 Homo sapi
6	734	99.2	770	6	AX118619	AX118619 Sequence
7	734	99.2	770	6	AF184073	AF184073 Homo sapi
8	734	99.2	782	6	AX459863	AX459863 Sequence
9	734	99.2	782	6	AB023416	AB023416 Homo sapi
10	733	99.1	779	6	AX017270	AX017270 Sequence
11	733	99.1	779	6	AX524974	AX524974 Sequence
12	733	99.1	779	6	BD134441	BD134441 Human nuc
13	730.8	98.8	785	6	BD057255	BD057255 Apoptosis
14	713	96.4	740	6	AK000211	AK000211 Homo sapi
15	687	92.8	712	9	AF310103	AF310103 Homo sapi
16	680	91.9	772	9	BC013569	BC013569 Homo sapi
17	610	82.4	713	6	AX118641	AX118641 Sequence
18	610	82.4	713	9	AF255794	AF255794 Homo sapi
19	585	79.1	585	6	AR256268	AR256268 Sequence
20	585	79.1	585	6	AX082248	AX082248 Sequence
21	585	79.1	585	6	AX536210	AX536210 Sequence
22	448	60.5	450	12	BT007571	BT007571 Synthetic
23	437.6	59.1	768	9	BC004470	BC004470 Homo sapi
24	416.8	56.3	588	4	AB050006	AB050006 Bos tauru
25	378	51.1	405	6	AX118643	AX118643 Sequence
26	360	48.6	2821	6	AX118618	AX118618 Sequence
27	360	48.6	2821	9	AF184072	AF184072 Homo sapi
28	360	48.6	12769	9	AC009088	AC009088 Homo sapi
29	356	48.1	356	6	AX118626	AX118626 Sequence
30	336	45.4	732	10	BC008252	BC008252 Mus muscu
31	336	45.4	770	10	AB032249	AB032249 Mus muscu
32	336	45.4	777	6	AR256273	AR256273 Sequence
33	336	45.4	777	6	AX082258	AX082258 Sequence
34	336	45.4	777	6	AX536205	AX536205 Sequence
35	336	45.4	777	6	AX536222	AX536222 Sequence
36	336	45.4	803	6	AX118637	AX118637 Sequence
37	333	45.0	701	10	AF310104	AF310104 Mus muscu
38	322	43.5	626	6	AX118621	AX118621 Sequence
39	321	43.4	340	6	AX118622	AX118622 Sequence
40	316.2	42.7	579	6	AR256274	AR256274 Sequence
41	316.2	42.7	579	6	AX082260	AX082260 Sequence
42	316	42.7	582	10	AB053165	AB053165 Rattus no
43	309.8	41.9	321	6	AX408419	AX408419 Sequence
44	293.8	39.7	579	6	AX536207	AX536207 Sequence
45	276.8	37.4	605	6	AX118639	AX118639 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR256267 740 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 48 from patent US 6482933.
ACCESSION AR256267
VERSION AR256267.1 GI:27305690
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 740)
AUTHORS Bertin,J.
TITLE Molecules of the card-related protein family and uses thereof
JOURNAL Patent: US 6482933-A 48 19-NOV-2002;
FEATURES Location/Qualifiers


```
QY 721 ATACGAAAAGCCAGCTTGAA 740
|||||
Db 721 ATACGAAAAGCCAGCTTGAA 740

RESULT 3
AX536208
LOCUS 740 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 4 from Patent WO0244354.
ACCESSION AX536208
VERSION AX536208.1 GI:25262599
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bertin,J.
TITLE Novel molecules of the card-related protein family and uses thereof
JOURNAL Patent: WO 0244354-A 4 06-JUN-2002;
MILLENIUM PHARMACEUTICALS, INC. (US)
FEATURES
source
location/Qualifiers
1..740
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
54..641
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57246.1"
/db_xref="GI:25262600"
/translation="MGRARDAITDALENLTAEELKFKLKLSVPLREYGRIPRGAL
LSMDALDITDKLVSVFLVTVGAEALTANVLRDMLQEMAGLOAATHQSGRAPAGIOA
PQSAKPGELHFDHRAALIARVTNVEWLLDALYKGLTDEYQAVRAEPTNPMSMR
KLFSTPAWNWTKLULLQALRESYLVEDLERS"
BASE COUNT 146 a 238 c 236 g 120 t
ORIGIN
Query Match 100.0%; Score 740; DB 6; Length 740;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGTGAGCGGGGAGCGCGGCGGATCCTGAGGCATGGGCG 60
Db 1 CGCGTCCGGCTGACGGGGTGAGCGGGGAGCGCGGCGGATCCTGAGGCATGGGCG 60

QY 61 GCGCGCGACGCCATCTCTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGT 120
Db 61 GCGCGCGACGCCATCTCTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGT 120

QY 121 TCAAGCTGAAGCTGTGTCGGTGCCGTCGCGAGGGCTACCGGCGCATCCCGCGGGCG 180
Db 121 TCAAGCTGAAGCTGTGTCGGTGCCGTCGCGAGGGCTACCGGCGCATCCCGCGGGCG 180

QY 181 CGCTGCTGTCATGAGCGCTTGAGACTCAACGACAAAGCTGTGTCAGCTTCTACCTGGAGA 240
Db 181 CGCTGCTGTCATGAGCGCTTGAGACTCAACGACAAAGCTGTGTCAGCTTCTACCTGGAGA 240

QY 241 CCTACGGCGCGAGCTCAACCGCTAAACGTCGCGACATGGGCCTGACGAGAGATGGCGG 300
Db 241 CCTACGGCGCGAGCTCAACCGCTAAACGTCGCGACATGGGCCTGACGAGAGATGGCGG 300

QY 301 GGCAGCTGACGGCGGACGACACAGGCTTGGAGCGGCGGCGAGCTGGGATCCAGGCC 360
Db 301 GGCAGCTGACGGCGGACGACACAGGCTTGGAGCGGCGGCGAGCTGGGATCCAGGCC 360

QY 361 CTCCTCAGTCGGCAGCAAGCCAGGCTGCACCTTTATAGACACAGCACCGGGCTGCCCTTA 420
Db 361 CTCCTCAGTCGGCAGCAAGCCAGGCTGCACCTTTATAGACACAGCACCGGGCTGCCCTTA 420

QY 421 TCGCAGGGGTCAAAACGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGGTCCTGACGG 480
|||||
```

Db 440 GGCGAGCTGCGAGCGCGCCAGCCAGCGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCC 381

Qy 361 CTCCTCAGTCGCGAGCCAGCCAGCGCTGCACCTTTATAGACCAGCACCGGCTGCGCTTA 420

Db 380 CTCCTCAGTCGCGAGCCAGCCAGCGGCTGCACCTTTATAGACCAGCACCGGCTGCGCTTA 321

Qy 421 TCGGAGGGTCAAAACGTTGAGTGCTGTGATGCTCTGTACGGGAAGGCTCTGACGG 480

Db 320 TCGGAGGGTCAAAACGTTGAGTGCTGTGATGCTCTGTACGGGAAGGCTCTGACGG 261

Qy 481 ATGAGCAGTACAGGCGAGTCGCGCGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540

Db 260 ATGAGCAGTACAGGCGAGTCGCGCGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 201

Qy 541 TCAGTTTACACAGCGCTGGAACCTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGG 600

Db 200 TCAGTTTACACAGCGCTGGAACCTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGG 141

Qy 601 AGTCCAGTCTTACCTGGTGGAGGACTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660

Db 140 AGTCCAGTCTTACCTGGTGGAGGACTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 81

Qy 661 TCGGTCAGCGCCCTGGCAATCCACCAATCATCTGAAATCTGATCTTTTATACAAAT 720

Db 80 TCGGTCAGCGCCCTGGCAATCCACCAATCATCTGAAATCTGATCTTTTATACAAAT 21

Qy 721 ATACGAAAGCCAGCTTGAA 740

Db 20 ATACGAAAGCCAGCTTGAA 1

RESULT 5

AF384665 740 bp mRNA linear PRI 19-JUN-2001

LOCUS Homo sapiens caspase recruitment domain protein 5 mRNA, complete cds.

DEFINITION AF384665.1 GI:14488058

VERSION AF384665.1

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 740)

AUTHORS Bertin,J.

TITLE CARD5 Protein is a CARD/PYRIN family member that is involved in apoptosis signal transduction

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 740)

AUTHORS Bertin,J.

TITLE Direct Submission

JOURNAL Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA

FEATURES

Location/Qualifiers

1..740

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

54..641

/note="CARD5; CARD/PYRIN family member; bipartite protein comprised of an N-terminal PYRIN domain and a C-terminal CARD domain; involved in apoptosis signal transduction"

/codon_start=1

/product="caspase recruitment domain protein 5"

/protein_id="AAK63850.1"

/db_xref="GI:14488059"

/translation="MGRARDAJLDALENLTABELKFKLLKLVLPREGYGRIPRGAL LSMADLDTDLKLVSVLEFYTGAEITANVRMDGLQEMAGQLQAAHQSGRAPAQIA PQQSAKPGHLHIDQHRALIAIARVNVNVEMLDALYKGLVLDQYQAVRAEPTNPSKMR KLFSTFPANWTKDLKLLQALRESQSIVLEDLERS"

146 a 238 c 236 g 120 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 740; DB 9; Length 740;

Best Local Similarity 100.0%; Pred. No. 3.1e-107;

Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGTCGCGCTGACAGCGGGGTGAGCGCGCGAGCGCGCGGGATCCTGGAGCCATGGGGC 60

Db 1 CGCGTCGCGCTGACAGCGGGGTGAGCGCGCGAGCGCGCGGGATCCTGGAGCCATGGGGC 60

Qy 61 GCGCGCGGAGCGCCATCCTGGATGCGCTGGAGAACTGACCGCGCGAGGAGCTCAAGAGT 120

Db 61 GCGCGCGGAGCGCCATCCTGGATGCGCTGGAGAACTGACCGCGCGAGGAGCTCAAGAGT 120

Qy 121 TCAAGCTGAAGCTGTGTGCTGCGCTGCGCGAGGAGTACCGGCGCATCTCCCGCGGGCG 180

Db 121 TCAAGCTGAAGCTGTGTGCTGCGCTGCGCGAGGAGTACCGGCGCATCTCCCGCGGGCG 180

Qy 181 CGCTCTGTTCATGACCGCTTGGACCTTACCGCAAGCTGTGTACCTTCTACCTGGAGA 240

Db 181 CGCTCTGTTCATGACCGCTTGGACCTTACCGCAAGCTGTGTACCTTCTACCTGGAGA 240

Qy 241 CCTAGCGCGCGAGCTCACCGCTAACGTCGTGCGCGACATGGGCTGCGAGAGATGGCG 300

Db 241 CCTAGCGCGCGAGCTCACCGCTAACGTCGTGCGCGACATGGGCTGCGAGAGATGGCG 300

Qy 301 GGCAGCTGCGAGCGCGCCAGCACAGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCC 360

Db 301 GGCAGCTGCGAGCGCGCCAGCACAGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCC 360

Qy 361 CTCCTCAGTCGCGCAAGCCAGCGCTTATATAGACAGCACCGGGCTGCGCTTA 420

Db 361 CTCCTCAGTCGCGCAAGCCAGCGCTTATATAGACAGCACCGGGCTGCGCTTA 420

Qy 421 TCGCGAGGTCACAAACCTTGAGTGGCTGCTGAGTGTCTGTACGGGAGGCTCTGACGG 480

Db 421 TCGCGAGGTCACAAACCTTGAGTGGCTGCTGAGTGTCTGTACGGGAGGCTCTGACGG 480

Qy 481 ATGAGCAGTACAGGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

Db 481 ATGAGCAGTACAGGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

Qy 541 TCAGTTTACACAGCGCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAAC 600

Db 541 TCAGTTTACACAGCGCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAAC 600

Qy 601 AGTCCAGTCTTACCTGCTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660

Db 601 AGTCCAGTCTTACCTGCTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660

Qy 661 TCCGGTCAGCGCCCTGGCAATCCACCAATCATCTGAAATCTGATCTTTTATACAAAT 720

Db 661 TCCGGTCAGCGCCCTGGCAATCCACCAATCATCTGAAATCTGATCTTTTATACAAAT 720

Qy 721 ATACGAAAGCCAGCTTGAA 740

Db 721 ATACGAAAGCCAGCTTGAA 740

RESULT 6

AX118619 770 bp DNA linear PAT 11-MAY-2001

LOCUS AX118619

DEFINITION Sequence 2 from Patent WO0129235.

ACCESSION AX118619

VERSION AX118619.1 GI:14035570

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Vertino,P.M.

TITLE Tms1 compositions and methods of use

JOURNAL Patent: WO 0129235-A 2 26-APR-2001; Emory University (US)

FEATURES		Location/Qualifiers	
source	RESULT 7	1..770	
	AF184073	/organism="Homo sapiens"	
	LOCUS	/mol_type="genomic DNA"	
	DEFINITION	/db_xref="taxon:9606"	
CDS	AF184073	75..662	
	AF184073.1	/note="unnamed protein product"	
	AF184073	/codon_start=1	
	AF184073.1	/protein_id="CAC38471.1"	
BASE COUNT	157 a	247 c	241 g
	125 t		
	ORIGIN		
	Query Match	99.2%;	Score 734; DB 6; Length 770;
Matches	Best Local Similarity	100.0%;	Pred. No. 2.8e-106;
	Matches 734;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
	QY	7	CGGCTGCAGCGGGTGCAGCGCGCGCGGATCCTGAGCATGGCGCGCGC 66
	Db	28	CGGCTGCAGCGGGTGCAGCGCGCGCGGATCCTGAGCATGGCGCGCGC 87
QY	67	GGAGCGCCATCCTGGATGCGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAAAGTTCAAGC 126	
	Db	88	GGAGCGCCATCCTGGATGCGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAAAGTTCAAGC 147
QY	127	TGAAGTGTGTGCGGCGCGCTGCGGAGGGCTACGGGGCGATCCCGGGGGCGGCTGC 186	
	Db	148	TGAAGTGTGTGCGGCGCGCTGCGGAGGGCTACGGGGCGATCCCGGGGGCGGCTGC 207
QY	187	TGTCATGACCGCTTGACCTCACGCAAGCTGTGACCTTACCTGGAGACCTAGC 246	
	Db	208	TGTCATGACCGCTTGACCTCACGCAAGCTGTGACCTTACCTGGAGACCTAGC 267
QY	247	GGCGCCAGCTCACCGCTTAACGTGCTCGCGGACATGGCGCTGCAGGAGATGGCGGCGAGC 306	
	Db	268	GGCGCCAGCTCACCGCTTAACGTGCTCGCGGACATGGCGCTGCAGGAGATGGCGGCGAGC 327
QY	307	TGAGCGCGCCACGACACCGAGGCTGTGAGCGCGCGCAGCTGGATTCAGGCGCCCTCTTC 366	
	Db	328	TGAGCGCGCCACGACACCGAGGCTGTGAGCGCGCGCAGCTGGATTCAGGCGCCCTCTTC 387
QY	367	AGTCGCGCAGCCAGCCAGCGCTGCACTTTATAGACGACACCGCGCTGCCTTATCGCA 426	
	Db	388	AGTCGCGCAGCCAGCCAGCGCTGCACTTTATAGACGACACCGCGCTGCCTTATCGCA 447
QY	427	GGGTCAACAAAGTTGAGTGGCTGCTGGATGCTGTACGGGAAGTCTCTGACGGATGAGC 486	
	Db	448	GGGTCAACAAAGTTGAGTGGCTGCTGGATGCTGTACGGGAAGTCTCTGACGGATGAGC 507
QY	487	AGTACAGGCGAGTGGCGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAGTT 546	
	Db	508	AGTACAGGCGAGTGGCGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAGTT 567
QY	547	TCACACCGCTGGAAGTGCAGACTGCAAGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 606	
	Db	568	TCACACCGCTGGAAGTGCAGACTGCAAGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 627
QY	607	AGTCCTACCTGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666	
	Db	628	AGTCCTACCTGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 687
QY	667	CAGCCCTGCAATCCCAACAAATCATCTGAAATCTGATCTTTTATACAAATATACGA 726	
	Db	688	CAGCCCTGCAATCCCAACAAATCATCTGAAATCTGATCTTTTATACAAATATACGA 747
QY	727	AAAGCCAGCTTGAA 740	
	Db	748	AAAGCCAGCTTGAA 761

AF184073	770 bp	mRNA	linear	PRI 02-FEB-2001
Homo sapiens target of methylation-induced silencing 1 (TMS1) mRNA, complete cds.				
AF184073				
AF184073.1	GI:9863863			
Homo sapiens (human)				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
1 (bases 1 to 770)				
Conway, K.E., McConnell, B.B., Bowring, C.E., Donald, C.D., Warren, S.T. and Vertino, P.M.				
TMS1, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers				
Cancer Res.	60 (22),	6236-6242	(2000)	
2052139				
11103776				
2 (bases 1 to 770)				
McConnell, B.B. and Vertino, P.M.				
Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1				
Cancer Res.	60 (22),	6243-6247	(2000)	
2052140				
11103777				
3 (bases 1 to 770)				
Vertino, P.M.				
Direct Submission				
Submitted (09-SEP-1999) Radiation Oncology, Emory University School of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA				
Location/Qualifiers				
1..770				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="IMAGE:159538; IMAGE:1990931"				
/clone_lib="Soares breast 3NBH8et; NCI_CGAP_Ut3 uterine tumor"				
1..770				
/gene="TMS1"				
75..662				
/gene="TMS1"				
/function="involved in apoptosis"				
/note="contains caspase-recruitment domain; CARD protein; target of methylation-mediated gene silencing in Homo sapien tumors"				
/codon_start=1				
/product="target of methylation-induced silencing 1"				
/protein_id="AAG01188.1"				
/db_xref="GI:9863864"				
/translation="MGRARDAILDALNLTAEELKKFKLLSVPLREGYGRIPRGAL LSMDALDTDKLVSYLETYGAEITANVLRMDLQEMAGLQAOATHQSSGAAPAGIQAPQSAKPLGHPIDHRAALIAITVNVEWLLDALYGVLTDEYQAVRAEPTNPSSMR KLFSTPAWNWTCDDLLQLALRESOSYLVEDLERS"				
BASE COUNT	157 a	247 c	241 g	125 t
ORIGIN				
Query Match				
Best Local Similarity				
Matches 734; Conservative				
0; Mismatches 0; Indels 0; Gaps 0;				
QY	7	CGGCTGCAGCGGGTGCAGCGCGCGGATCCTGAGCATGGCGCGCGC 66		
Db	28	CGGCTGCAGCGGGTGCAGCGCGCGGATCCTGAGCATGGCGCGCGC 87		
QY	67	GGAGCGCCATCCTGGATGCGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAAAGTTCAAGC 126		
Db	88	GGAGCGCCATCCTGGATGCGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAAAGTTCAAGC 147		

QY	127	TGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTCC	186
Db	148	TGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTCC	207
QY	187	TGTCCATGAGCGCTTGGACCTCACCAGCAAGCTGTGAGCTTCTACCTGGAGACCTAGC	246
Db	208	TGTCCATGAGCGCTTGGACCTCACCAGCAAGCTGTGAGCTTCTACCTGGAGACCTAGC	267
QY	247	GGCGCGAGCTCACCCTTAACGCTGCTGCGGACATGGGCTGAGGAGATGGCGGCGCAGC	306
Db	268	GGCGCGAGCTCACCCTTAACGCTGCTGCGGACATGGGCTGAGGAGATGGCGGCGCAGC	327
QY	307	TGAGCGCGCCAGCAGCGGCTTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCCCTC	366
Db	328	TGAGCGCGCCAGCAGCGGCTTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCCCTC	387
QY	367	AGTCGCGAGCAAGCAGCGCTGCACTTTATAGACCAGCAGCGGCTGGGATCCAGGCTCC	426
Db	388	AGTCGCGAGCAAGCAGCGCTGCACTTTATAGACCAGCAGCGGCTGGGATCCAGGCTCC	447
QY	427	GGGTCAACAACGTTGAGTGGCTGCTGATGCTGTACGGGAGGTCCTGACGGATGAGC	486
Db	448	GGGTCAACAACGTTGAGTGGCTGCTGATGCTGTACGGGAGGTCCTGACGGATGAGC	507
QY	487	AGTACCGAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	546
Db	508	AGTACCGAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	567
QY	547	TCACACGCTGGAACCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	606
Db	568	TCACACGCTGGAACCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	627
QY	607	AGTCCTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGTT	666
Db	628	AGTCCTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGTT	687
QY	667	CAGCCCTGGCAATCCCAACATCATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	726
Db	688	CAGCCCTGGCAATCCCAACATCATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	747
QY	727	AAAGCCAGCTTGAA	740
Db	748	AAAGCCAGCTTGAA	761

RESULT 8
 AX459863
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Tschoopp,J. and Martinon,F.
 Proteins and dna sequences underlying these proteins used for
 treating inflammations
 Patent: WO 0240668-A 6 23-MAY-2002;
 Apotech Research and Development Ltd. (CH)
 Location/Qualifiers
 1..782
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Pycard.cdna"
 158 a 251 c 246 g 127 t
 BASE COUNT
 ORIGIN
 Query Match 99.2%; Score 734; DB 6; Length 782;
 Best Local Similarity 100.0%; Pred. No. 2.8e-106;

Matches	734;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	7	CGGCTGACGCGGGTACGCGCGGCGGAGTCTTGAGCCATGGGCGCGCGC	66						
Db	40	CGGCTGACGCGGGTACGCGCGGCGGAGTCTTGAGCCATGGGCGCGCGC	99						
QY	67	GCGACCGCATCTGGATGCGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC	126						
Db	100	GCGACCGCATCTGGATGCGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC	159						
QY	127	TGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	186						
Db	160	TGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	219						
QY	187	TGTCCATGAGCGCTTGGACCTCACCAGCAAGCTGTGAGCTTCTACCTGGAGACCTACG	246						
Db	220	TGTCCATGAGCGCTTGGACCTCACCAGCAAGCTGTGAGCTTCTACCTGGAGACCTACG	279						
QY	247	GCGCGAGCTCACCCTTAACGCTGCTGCGGACATGGGCTGCGGAGATGGCGGCGCAGC	306						
Db	280	GCGCGAGCTCACCCTTAACGCTGCTGCGGACATGGGCTGCGGAGATGGCGGCGCAGC	339						
QY	307	TGACGCGCGCCAGCAGCGGCTTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCCCTC	366						
Db	340	TGACGCGCGCCAGCAGCGGCTTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCCCTC	399						
QY	367	AGTCGCGAGCAAGCAGCGCTGCACTTTATAGACCAGCAGCGGCTGCGCTTATCGGA	426						
Db	400	AGTCGCGAGCAAGCAGCGCTGCACTTTATAGACCAGCAGCGGCTGCGCTTATCGGA	459						
QY	427	GCGTCAACAACGTTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	486						
Db	460	GCGTCAACAACGTTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	519						
QY	487	AGTACGAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	546						
Db	520	AGTACGAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	579						
QY	547	TCACACGAGCTGGAACCTGCAAGGACTTGCTCTCCAGAGCCCTCAAGGAGTCTCC	606						
Db	580	TCACACGAGCTGGAACCTGCAAGGACTTGCTCTCCAGAGCCCTCAAGGAGTCTCC	639						
QY	607	AGTCCTACCTGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGTT	666						
Db	640	AGTCCTACCTGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGTT	699						
QY	667	CAGCCCTGGCAATCCCAACATCATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	726						
Db	700	CAGCCCTGGCAATCCCAACATCATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT	759						
QY	727	AAAGCCAGCTTGAA	740						
Db	760	AAAGCCAGCTTGAA	773						
RESULT 9									
AB023416									
LOCUS									
DEFINITION									
Homo sapiens ASC mRNA for apoptosis-associated speck-like protein containing a CARD, complete cds.									
AB023416									
ACCESSION									
VERSION									
KEYWORDS									
SOURCE									
ORGANISM									
Homo sapiens (human)									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
1 (sites)									
REFERENCE									
AUTHORS									
TITLE									
J. Biol. Chem. 274 (48), 33835-33838 (1999)									

AB023416
 Homo sapiens
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Masumoto,J., Taniguchi,S., Ayukawa,K., Sarvotham,H., Kishino,T.,
 Niiikawa,N., Hidaka,E., Katsuyama,T., Higuchi,T. and Sagara,J.
 ASC, a novel 22-kDa protein, aggregates during apoptosis of human
 promyelocytic leukemia HL-60 cells
 J. Biol. Chem. 274 (48), 33835-33838 (1999)

QY	367	AGTGGCAGCCAGCCAGGCGCTGCACTTTATAGACAGCACCAGGCTGGGCTGGCTTTATCGCGA	426
Db	407	AGTGGCAGCCAGCCAGGCGCTGCACTTTATAGACAGCACCAGGCTGGCTTTATCGCGA	466
QY	427	GGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGAGATGAGC	486
Db	467	GGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGAGATGAGC	526
QY	487	AGTACAGCAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	546
Db	527	AGTACAGCAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	586
QY	547	TCACACAGCCTGGAATCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC	606
Db	587	TCACACAGCCTGGAATCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC	646
QY	607	AGTCTTACTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT	666
Db	647	AGTCTTACTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT	706
QY	667	CAGCCCTGGCAATCCCAATATCATCTGATCTGATCTTTTATACAAATATACGA	726
Db	707	CAGCCCTGGCAATCCCAATATCATCTGATCTTTTATACAAATATACGA	766
QY	727	AAAGCCAGCTTGA	739
Db	767	AAAGCCAGCTTGA	779
RESULT 11			
LOCUS	AX524974	779 bp . DNA	linear PAT 21-NOV-2002
DEFINITION	Sequence 21 from Patent EP1236799.		
ACCESSION	AX524974		
VERSION	AX524974.1 GI:25170056		
KEYWORDS	Human sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Specht, T., Hinzmann, B., Schmitt, A., Pilarsky, C., Dahl, E. and Rosenthal, A.		
TITLE	Human nucleic acid sequences derived from breast tumor tissue		
JOURNAL	Patent: EP 1236799-A 21 04-SEP-2002;		
metagen Pharmaceuticals GmbH (DE)			
FEATURES	Location/Qualifiers		
source	1..779		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	149 a 252 c 251 g 127 t		
ORIGIN			
Query Match	99.1%; Score 733; DB 6; Length 779;		
Best Local Similarity	100.0%; Pred. No. 4e-106;		
Matches	733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	7	CGGCTCAGCGGGTGAGCGGCGGAGCGCGCGGGATCTGGAGCCATGGGCGCGCGC	66
Db	47	CGGCTCAGCGGGTGAGCGGCGGAGCGCGCGGGATCTGGAGCCATGGGCGCGCGC	106
QY	67	GGACGCCATCTGGATGGCTGGAGAACCTGACCCCGAGGAGCTCAAGAAGTTCAAGC	126
Db	107	GGACGCCATCTGGATGGCTGGAGAACCTGACCCCGAGGAGCTCAAGAAGTTCAAGC	166
QY	127	TGAAGCTGTCTGGTGGCTGGAGGCTTACGGCGCATCCCGCGGCGCGCTGC	186
Db	167	TGAAGCTGTCTGGTGGCTGGAGGCTTACGGCGCATCCCGCGGCGCGCTGC	226
QY	187	TGTCCATGAGCGCTTGGACCTCACCGACAAGCTGGTCTACCTGGAGACTTACG	246
Db			

FEATURES

FT Key Location/Qualifiers
FT source 1..779 /organism="Homo sapiens (human)".
FT Location/Qualifiers

Db	227	TGTCCATGAGCGCTTGGACCTCACCGACAAGCTGGTCACTTTACCTGGAGACTACG	286
QY	247	GGCGCAGCTCACCGCTAACGCTGCGGCATGCGGCTGCGAGAGATGGCCGGGCGAGC	306
Db	287	GGCGCAGCTCACCGCTAACGCTGCGGCATGCGGCTGCGAGAGATGGCCGGGCGAGC	346
QY	307	TGCAGCGGCCACGCAACAGGGCTCTGGAGCGGCCAGCTGGGATCCAGGCCCTCTCTC	366
Db	347	TGCAGCGGCCACGCAACAGGGCTCTGGAGCGGCCAGCTGGGATCCAGGCCCTCTCTC	406
QY	367	AGTGGGAGCCAGCAGCGGCTGCACTTTATAGACAGCACCAGGCTGGCTTTATCGCGA	426
Db	407	AGTGGGAGCCAGCAGCGGCTGCACTTTATAGACAGCACCAGGCTGGCTTTATCGCGA	466
QY	427	GGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGAGATGAGC	486
Db	467	GGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGAGATGAGC	526
QY	487	AGTACAGCAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	546
Db	527	AGTACAGCAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAGTT	586
QY	547	TCACACAGCCTGGAATCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC	606
Db	587	TCACACAGCCTGGAATCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC	646
QY	607	AGTCTTACTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT	666
Db	647	AGTCTTACTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT	706
QY	667	CAGCCCTGGCAATCCCAATATCATCTGATCTTTTATACAAATATACGA	726
Db	707	CAGCCCTGGCAATCCCAATATCATCTGATCTTTTATACAAATATACGA	766
QY	727	AAAGCCAGCTTGA	739
Db	767	AAAGCCAGCTTGA	779
RESULT 12			
LOCUS	BD134441	779 bp DNA	linear PAT 18-SEP-2002
DEFINITION	Human nucleic acid sequence originating in mammary tumor tissue.		
ACCESSION	BD134441		
VERSION	BD134441.1 GI:23229386		
KEYWORDS	JP 2002506643-A/19.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 779)		
AUTHORS	Specht, T., Hintzman, B., Armin, S., Pirarski, C., Edgar, D. and Rosenthal, A.		
TITLE	Human nucleic acid sequence originating in mammary tumor tissue		
JOURNAL	Patent: JP 2002506643-A 19 05-MAR-2002;		
COMMENT	METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH		
	OS Homo sapiens (human)		
	PN JP 2002506643-A/19		
	PD 05-MAR-2002		
	PF 19-MAR-1999 JP 2000536852		
	PR 20-MAR-1998 DE 198 13 839.3		
	PI THOMAS SPEFT, BERND HINTZMAN, SHCMITT ARMIN, CHRISTIAN PIRARSKI, PI DUHL EDGAR, PI		
	PI ANDRE ROSENTHAL		
	PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P35/00, C07K14/47,		
	PC C07K16/18,		
	PC C12N1/19, C12N5/10, C12N15/00, A61K37/02, C12N5/00 CC		
	nucleic acid sequence originating in mammary tumor CC		
	tissue		
	FT Key Location/Qualifiers		
	FT source 1..779 /organism="Homo sapiens (human)".		
	FT Location/Qualifiers		

source 1..779
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606" 127 t

BASE COUNT 149 a 252 c 251 g 127 t

ORIGIN

Query Match 99.1%; Score 733; DB 6; Length 779;
Best Local Similarity 100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGCAGCGGGTGGAGCGCGGAGCGCGGGATCTCTGAGCCATGGGGCGGCGC 66
DB 47 CGGCTGCAGCGGGTGGAGCGCGGAGCGCGGGATCTCTGAGCCATGGGGCGGCGC 106
QY 67 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGC 126
DB 107 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGC 166
QY 127 TGAAGCTGCTGCTGGTGGCTGCGCGAGGGGTACGGGCGCATCCCGGGGCGGCGCTGC 186
DB 167 TGAAGCTGCTGCTGGTGGCTGCGCGAGGGGTACGGGCGCATCCCGGGGCGGCGCTGC 226
QY 187 TGTTCATGACGCGCTTGGACCTTACCGACAGCTGGTCAAGTCTTACCTGGAGACTTACG 246
DB 227 TGTTCATGACGCGCTTGGACCTTACCGACAGCTGGTCAAGTCTTACCTGGAGACTTACG 286
QY 247 GCGCGAGCTCACCGTCAAGTCTGCGGAGCATATGGGCGCTGAGGAGATGGCGGCGAGC 306
DB 287 GCGCGAGCTCACCGTCAAGTCTGCGGAGCATATGGGCGCTGAGGAGATGGCGGCGAGC 346
QY 307 TGCAGCGCGCCACGACCAAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCTC 366
DB 347 TGCAGCGCGCCACGACCAAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCTC 406
QY 367 AGTCGCGAGCCAGCCAGCGCTGCACTTTATAGACAGACCGCGGCTGGCTTTATCGCGA 426
DB 407 AGTCGCGAGCCAGCCAGCGCTGCACTTTATAGACAGACCGCGGCTGGCTTTATCGCGA 466
QY 427 GGGTCACAAAGCTTGAAGTGGCTGGATGCTCTGTACGGGAGGTCCTGACGGATGAGC 486
DB 467 GGGTCACAAAGCTTGAAGTGGCTGGATGCTCTGTACGGGAGGTCCTGACGGATGAGC 526
QY 487 AGTACCGAGCAGTGGCGGCGAGCCACCAAGCAAGATGCGGAGCTCTTCAGTT 546
DB 527 AGTACCGAGCAGTGGCGGCGAGCCACCAAGCAAGATGCGGAGCTCTTCAGTT 586
QY 547 TCACACGAGCTGGAACTGGACTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGGAGTCCC 606
DB 587 TCACACGAGCTGGAACTGGACTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGGAGTCCC 646
QY 607 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
DB 647 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 706
QY 667 CAGCCCTTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
DB 707 CAGCCCTTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 766
QY 727 AAAGCCAGCTTGA 739
DB 767 AAAGCCAGCTTGA 779

RESULT 13
BD057255
LOCUS 785 bp DNA linear PAT 27-AUG-2002
DEFINITION Apoptosis-related protein, its antibody and its DNA.
ACCESSION BD057255
VERSION BD057255.1 GI:22602861
KEYWORDS JP 2001275681-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 785)
Masumoto,J., Sagara,J. and Taniguchi,S.
Apoptosis-related protein, its antibody and its DNA
Patent: JP 2001275681-A 1 09-OCT-2001;
JUNYA MASUMOTO,JUNJI ATRA,SHUNICHIRO TANIGUCHI, MEDICAL &
BIOLOGICAL LABORATORIES CO LTD
OS Homo sapiens (human)
PN JP 2001275681-A/1
PD 09-OCT-2001
PF 31-MAR-2000 JP 2000098204
PI JUNYA MASUMOTO,JUNJI SAGARA,SHUNICHIRO TANIGUCHI PC
C12N15/09,C07K16/82,C07K16/32,C12P21/02,C12P21/08,C12N15/00 CC
FH Key Location/Qualifiers.
Location/Qualifiers
1..785
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 160 a 252 c 247 g 126 t
ORIGIN

Query Match 98.8%; Score 730.8; DB 6; Length 785;
Best Local Similarity 99.7%; Pred. No. 8.8e-106;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 CGGCTGCAGCGGGTGGAGCGCGGAGGATCTCTGAGCCATGGGGCGGCGC 66
DB 40 CGGCTGCAGCGGGTGGAGCGCGGAGGATCTCTGAGCCATGGGGCGGCGC 99
QY 67 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGC 126
DB 100 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGC 159
QY 127 TGAAGCTGCTGCTGGTGGCTGCGGAGGGGTACGGGCGCATCCCGGGGCGGCGCTGC 186
DB 160 TGAAGCTGCTGCTGGTGGCTGCGGAGGGGTACGGGCGCATCCCGGGGCGGCGCTGC 219
QY 187 TGTTCATGACGCGCTTGGACCTTACCGACAGCTGGTCAAGTCTTACCTGGAGACTTACG 246
DB 220 TGTTCATGACGCGCTTGGACCTTACCGACAGCTGGTCAAGTCTTACCTGGAGACTTACG 279
QY 247 GCGCGAGCTCACCGCTTAACGCTGTCGCGACATGGGCTTCAGGAGATGGCGGCGAGC 306
DB 280 GCGCGAGCTCACCGCTTAACGCTGTCGCGACATGGGCTTCAGGAGATGGCGGCGAGC 339
QY 307 TGCAGGCGGCACGACACAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCTC 366
DB 340 TGCAGGCGGCACGACACAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCTC 399
QY 367 AGTCGCGAGCCAAAGCCAGCGCTTATAGACAGCAGCGGCTGGCTTATCGCGA 426
DB 400 AGTCGCGAGCCAAAGCCAGCGCTTATAGACAGCAGCGGCTGGCTTATCGCGA 459
QY 427 GGGTCACAAAGCTTGAAGTGGCTGCTGATGCTCTGTACGGGAGGTCCTGACGAGTACG 486
DB 460 GGGTCACAAAGCTTGAAGTGGCTGCTGATGCTCTGTACGGGAGGTCCTGACGAGTACG 519
QY 487 AGTACCGAGCAGTGGGCGGAGCCACCAAGCAAGATGCGGAGCTTCTTCAGTT 546
DB 520 AGTACCGAGCAGTGGGCGGAGCCACCAAGCAAGATGCGGAGCTTCTTCAGTT 579
QY 547 TCACACCGAGCTGGAACTGGAGCTTCAAGGACTTCTCTCCAGGCGCTTAAGGGAGTCCC 606
DB 580 TCACACCGAGCTGGAACTGGAGCTTCAAGGACTTCTCTCCAGGCGCTTAAGGGAGTCCC 639
QY 607 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
DB 640 AGTCCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 699
QY 667 CAGCCCTTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726

```

Db      700 CAGCCCTCGCAATCCCAACAAATCATCTGAATCTGATCTTTTATACACAATATACGA 759
QY      727 AAAGCCAGCTTGAA 740
Db      760 AAAGCCAGCTTGAA 773

RESULT 14
AK000211
LOCUS   Homo sapiens cDNA FLJ20204 fis, clone COLF1505.
DEFINITION
ACCESSION AK000211
VERSION   AK000211.1 GI:7020146
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
          Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
          Nakamura,Y., Isogai,T. and Sugano,S.
          NEDO human cDNA sequencing project
          Unpublished
          2 (bases 1 to 740)
          Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
          Shibahara,T., Tanaka,T. and Nakamura,Y.
          Direct Submission
          Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1,
          Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ms.u-tokyo.ac.jp,
          Tel:81-3-5449-5286, Fax:81-3-5449-5416)
          NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing; Research Association for Biotechnology; cDNA library
          construction, 5'- & 3'-end one pass sequencing; Department of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
          Location/Qualifiers
          1..740
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="COLF1505"
             /tissue_type="colon mucosa"
             /clone_lib="Colf"
             /note="cloning vector pME18SFL3"
             177..626
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="BAA91012.1"
             /db_xref="GI:7020147"
             /translation="MDALDLDKLVSYFLETYGABELTANVLRMDGLQEMAGLQAAATH
             QSGAAPAGIQAPPOSAAPKPLHFDIQHRAALIAIARVNVWLLDLYGKVLITDEYQA
             VRAEPTNPMKMKLFSFTPAWNTCKDLILQALRESQSYLVEDLERS"
BASE COUNT 161 a 231 c 230 g 118 t
ORIGIN

Query Match 96.4%; Score 713; DB 9; Length 740;
Best Local Similarity 99.7%; Pred. No. 5.7e-103;
Matches 725; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY      14 ACGGGGTAGCGCGCGAGCGCGCGGGATCTCTGGAGCCATGGGCGCGCGGACGC 73
Db      1 ACGGGGTAGCGCGCGAGCGCGCGGGATCTCTGGAGCCATGGG--GGCGCGGACGC 58

QY      74 CATCTCTGGATCGCTGAGAACCTGACCGCCGAGGAGCTCAAGAGTCTCAAGCT 133
Db      59 CATCTCTGGATCGCTGAGAACCTGACCGCCGAGGAGCTCAAGAGTCTCAAGCT 118

QY      134 GCTGTCTGCTGCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTCTGTCCAT 193

```

```

Db      119 GCTGTCTGCTGCGCTGCGGAGGGGCTACGGGCGCATCCCGGGGGCGGCTGTCTCCAT 178
QY      194 GGACGCTTGGACCTTCAACCGACCAAGCTGGTCAAGCTTCTTACCTTGGAGAGCTTACGGGCGCGGA 253
Db      179 GGACGCTTGGACCTTCAACCGACCAAGCTGGTCAAGCTTCTTACCTTGGAGAGCTTACGGGCGCGGA 238
QY      254 GCTACCCCTAAACGCTGCGCGACATGGGCTGTCAGAGATGGCCGGGAGCTTCAGGCG 313
Db      239 GCTACCCCTAAACGCTGCGCGACATGGGCTGTCAGAGATGGCCGGGAGCTTCAGGCG 298
QY      314 GGCCACGACACAGGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTCTCTCAGTCCGC 373
Db      299 GGCCACGACACAGGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTCTCTCAGTCCGC 358
QY      374 AGCCAAAGCAGGCTTATAGACAGCAGCAGCGGCTGCGCTTATCGGAGGGTCCAC 433
Db      359 AGCCAAAGCAGGCTTATAGACAGCAGCAGCGGCTGCGCTTATCGGAGGGTCCAC 418
QY      434 AAACGTTGAGTGGCTTGGATGCTCTCTACGGGAAGTCTCTGACGGATGAGCAGTACCA 493
Db      419 AAACGTTGAGTGGCTTGGATGCTCTCTACGGGAAGTCTCTGACGGATGAGCAGTACCA 478
QY      494 GCGAGTGGCGCGGAGCCCAACACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACC 553
Db      479 GCGAGTGGCGCGGAGCCCAACACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACC 538
QY      554 AGCCTGGAACCTGGACCTGCAAGGACTTCTCTCCAGGCCCTTAAGGGAGTCCCACTCTTA 613
Db      539 AGCCTGGAACCTGGACCTGCAAGGACTTCTCTCCAGGCCCTTAAGGGAGTCCCACTCTTA 598
QY      614 CTTGTGGAGGACCTTGGAGCGGAGCTGAGGCTCTTCTCCAGCAACACCTCCGGTCAAGCCC 673
Db      599 CTTGTGGAGGACCTTGGAGCGGAGCTGAGGCTCTTCTCCAGCAACACCTCCGGTCAAGCCC 658
QY      674 TGGCAATCCCAACAAATCATCTGATCTGATCTTTTATACAAATATACAAAAGCCA 733
Db      659 TGGCAATCCCAACAAATCATCTGATCTGATCTTTTATACAAATATACAAAAGCCA 718
QY      734 GCTTGAA 740
Db      719 GCTTGAA 725

RESULT 15
AF310103
LOCUS   Homo sapiens PYCARD mRNA, complete cds.
DEFINITION
ACCESSION AF310103
VERSION   AF310103.1 GI:11096298
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
          The pyrin domain: a possible member of the death domain-fold family
          implicated in apoptosis and inflammation
          Curr. Biol. 11 (4), R118-R120 (2001)
JOURNAL   MEDLINE
PUBMED    11250163
REFERENCE
AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
          Direct Submission
          Submitted (28-SEP-2000) Institute of Biochemistry, University of
          Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
JOURNAL
FEATURES
source
1..712
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
1..588
   /note="contains PYD and CARD domains"
CDS

```

```
/codon_start=1
/product="PYCARD"
/protein_id="AAG30286.1"
/db_xref="GI:11096299"
/translation="MGRARDAILDALENLTAEELKFKLKLLSVPLREGYGRIPRGAL
LSMDALDITDKLVSYLETYNELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIOA
PPQSAKFLHFIIDHRAALFARVTNVEWLLDALYKVLITDEQYQAVRAEPTNPKMR
KLFSFPAWNTCKDLQLALRESQSYLVEDLERS"
BASE COUNT      166 a      221 c      210 g      115 t
ORIGIN
Query Match      92.8%; Score 687; DB 9; Length 712;
Best Local Similarity 100.0%; Pred. No. 7.2e-99;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 54 ATGGGGCGCGCGCGAGCCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTC 113
Db 1 ATGGGGCGCGCGCGAGCCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTC 60
QY 114 AAGAAATTCAAGCTGAAGTGTGTGCTGCGTGCCTGCGGAGGGCTACGGGCGCATCCCG 173
Db 61 AAGAAATTCAAGCTGAAGTGTGTGCTGCGTGCCTGCGGAGGGCTACGGGCGCATCCCG 120
QY 174 CGGGCGCGCTCTGTTCATGAGACGCTTGGACCTCACCGACAAGCTGAGTCTTCTAC 233
Db 121 CGGGCGCGCTCTGTTCATGAGACGCTTGGACCTCACCGACAAGCTGAGTCTTCTAC 180
QY 234 CTGGAGACCTACGGCGCGAGCTCACCGCTACCGTGTGCGGAGGCTGCGGCTGCAGGAG 293
Db 181 CTGGAGACCTACGGCGCGAGCTCACCGTGTGCGGAGGCTGCGGCTGCAGGAG 240
QY 294 ATGGCGGGGAGCTGCAGCGCGCCACGACACAGGGCTCTGGAGCGCGCGAGCTGGGATC 353
Db 241 ATGGCGGGGAGCTGCAGCGCGCCACGACACAGGGCTCTGGAGCGCGCGAGCTGGGATC 300
QY 354 CAGGCCCTCTCTCAGTGGCGAGCCAGCCAGCGCTGCACTTTATAGACAGCACCGGGCT 413
Db 301 CAGGCCCTCTCTCAGTGGCGAGCCAGCCAGCGCTGCACTTTATAGACAGCACCGGGCT 360
QY 414 GCGCTTATCGCGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGTC 473
Db 361 GCGCTTATCGCGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGTC 420
QY 474 CTGACGGATGAGCAGTACACAGCAGTGCAGGGCGGAGCCCAACCAAGCAAGATGCGG 533
Db 421 CTGACGGATGAGCAGTACACAGCAGTGCAGGGCGGAGCCCAACCAAGCAAGATGCGG 480
QY 534 AAGCTTTCAGTTTCACACAGCCTGGAACTGGACCTGCAAGACTTGCTCTCCAGGCC 593
Db 481 AAGCTTTCAGTTTCACACAGCCTGGAACTGGACCTGCAAGACTTGCTCTCCAGGCC 540
QY 594 CTAGGGAGTCCAGTCTCTACCTGGTGGAGGACCTGGAGGGAGCTGAGGCTCTTCCCA 653
Db 541 CTAGGGAGTCCAGTCTCTACCTGGTGGAGGACCTGGAGGGAGCTGAGGCTCTTCCCA 600
QY 654 GCAACACTCCGGTACGCCCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTAT 713
Db 601 GCAACACTCCGGTACGCCCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTAT 660
QY 714 ACACAATATACGAAAGCCAGCTTGAA 740
Db 661 ACACAATATACGAAAGCCAGCTTGAA 687
```

Search completed: January 29, 2004, 09:40:36
Job time : 2665.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:06 ; Search time 182.487 Seconds
(without alignments)
10946.442 Million cell updates/sec

Title: US-09-996-617-7

Perfect score: 740

Sequence: 1 cgcgtccggctgcagcggg.....atcgaaagccagcttgaa 740

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	740	100.0	740	22	AAF30007 Human CARD-5 cDNA.
2	740	100.0	740	24	ABK87966 Human caspase recr
3	740	100.0	740	24	ABK87967 Human caspase recr
4	740	100.0	740	25	ABSS56032 cDNA encoding huma
5	734	99.2	770	22	AAD03890 Human target of me
6	734	99.2	782	24	AAL47126 Pyrin domain conta
7	733	99.1	745	22	AAAS59817 Human novel cytoki
8	733	99.1	779	20	AAZ33631 Human breast tumou

9	721.2	97.5	806	22	AAH34052 Human colon cancer
10	697.6	94.3	811	21	AAAC77884 Human cancer assoc
11	610	82.4	713	22	AAD03906 Alternatively spli
12	531	71.8	639	22	AAH34053 Human colon cancer
13	384.6	52.0	432	21	AAC98638 Human colon cancer
14	378	51.1	405	22	AAD03907 Alternatively spli
15	360	48.6	2821	22	AAD03889 Human target of me
16	356	48.1	356	22	AAD03894 Human target of me
17	336	45.4	777	22	AAF30008 Mouse CARD-5 cDNA.
18	336	45.4	777	24	ABK87964 Mouse caspase recr
19	336	45.4	777	24	ABK87965 Mouse caspase recr
20	336	45.4	803	22	AAD03904 Mouse target of me
21	322	43.5	626	22	AAD03891 CpG island of huma
22	321	43.4	340	22	AAD03892 Human target of me
23	309.8	41.9	321	24	ABN94568 Gene #1066 used to
24	276.8	37.4	605	22	AAD03905 Rat target of meth
25	177.6	24.0	551	24	AAL47124 Pyrin domain conta
26	174.4	23.6	807	22	AAC98099 Human FLEXHT-30 nu
27	115	15.5	4194	22	AAD02761 Human NB-ARC and C
28	115	15.5	4329	22	AAD02762 Human NB-ARC and C
29	115	15.5	4422	22	AAD02760 Pyrin domain conta
30	115	15.5	5100	24	AAL47127 Human G-protein co
31	115	15.5	5122	22	AAD12951 Human CARD-7 polyp
32	115	15.5	5523	22	AAF83651 cDNA encoding huma
33	115	15.5	5523	25	ABSS5497 cDNA encoding huma
34	115	15.5	5523	21	ABSS56030 Human secreted pro
35	111.4	15.1	2657	21	AAA78392 Human novel cytoki
36	98.6	13.3	180	22	AAAS59860 MEV gene sequence
37	75.6	10.2	16891	20	AAK37084 Pyrin domain conta
38	74.2	10.0	3500	24	AAL47125 FWP associated pro
39	74.2	10.0	3512	20	AAK37085 Human breast cell
40	74	10.0	394	22	ABA50929 Human foetal liver
41	74	10.0	394	22	ABA68906 Probe #14325 for g
42	74	10.0	394	22	ABA35859 Human brain expres
43	74	10.0	394	22	AAK17239 Human bone marrow
44	74	10.0	394	22	AAK43026 Probe #13724 for g
45	74	10.0	394	22	AAI23791

ALIGNMENTS

RESULT 1
AAF30007
ID AAF30007 standard; cDNA; 740 BP.
XX
AC AAF30007;
XX
XX 23-APR-2001 (first entry)
DT
DE Human CARD-5 cDNA.
XX
KW CARD-5; caspase recruitment domain; human; cancer; infection;
KW autoimmune disease; neurological disease; haematological disease;
KW immune disease; inflammation; antitumor; antiseptic;
KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW gene therapy; ss.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 54..641
CDS /*tag= a
FT /note= "the open reading frame is also specifically
FT claimed in Claim 1(a)"

XX (MILL-) MILLENNIUM PHARM INC.
PA Bertin J;
XX
XX
XX
XX WPI; 2001-061973/07.
DR P-PSDB; AAB20085.
XX
XX Isolated intracellular proteins predicted to be involved in regulating
PT caspase activation are used for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PT haematological disorders -
XX
XX Claim 1(a); Fig 21; 208pp; English.
XX
XX The present sequence is that of cDNA encoding human caspase
CC recruitment domain 5 (CARD-5, see AAB20085). The cDNA was isolated
CC from a testis cDNA library using murine CARD-L. Plasmid EPHC5
CC containing CARD-5 cDNA is deposited as ATCC PTA-213. CARD-5
CC is an intracellular protein predicted to be involved in regulating
CC caspase activation. It is useful as a modulating agent in
CC regulating cellular processes include cell growth and cell death.
CC Methods of diagnosing and treating patients suffering from a
CC disorder associated with an abnormal level or rate of apoptotic
CC cell death, abnormal activity of the Fas/APO-1 receptor complex,
CC abnormal activity of the tumour necrosis factor receptor complex
CC or abnormal activity of a caspase involve administering a compound
CC that modulates the expression or activity of CARD-3, CARD-4, CARD-5
CC or CARD-6 e.g. a small molecule, antisense nucleic acid, ribozyme
CC or polypeptide. Such disorders include cancer, viral infection,
CC autoimmune disorders, neurological diseases, haematological
CC disorders, inflammatory disorders and immune disorders. CARD
CC nucleic acids can be used to express CARD proteins in a host cell
CC e.g. for gene therapy applications, to detect a genetic lesion and
CC to modulate CARD activity.
XX
XX Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;
SQ

Query Match 100.0%; Score 740; DB 22; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.7e-147;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGGTGAGCGGGGAGCGGGGATCTGGAGGCATGGGGC 60
DB 1 CGCGTCCGGCTGACGGGGGTGAGCGGGGAGCGGGGATCTGGAGGCATGGGGC 60

QY 61 GCGCGCGACGCCATCTCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGT 120
DB 61 GCGCGCGACGCCATCTCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGT 120

QY 121 TCAAGCTGAAGTGTCTGCTGGTGGCTGGAGGAGGCTACGGGCGCATCCCGCGGGCG 180
DB 121 TCAAGCTGAAGTGTCTGCTGGTGGCTGGAGGAGGCTACGGGCGCATCCCGCGGGCG 180

QY 181 CGCTGCTGCTGATGACGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTGGAGA 240
DB 181 CGCTGCTGCTGATGACGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTGGAGA 240

QY 241 CCTACGGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTGACGAGATGGCGG 300
DB 241 CCTACGGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTGACGAGATGGCGG 300

QY 301 GCGAGCTGACGGGCGCAGCAGCAGGCTTGGAGCGCGCGCAGTGGGATCCAGGGCC 360
DB 301 GCGAGCTGACGGGCGCAGCAGCAGGCTTGGAGCGCGCGCAGTGGGATCCAGGGCC 360

QY 361 CTCCTCAGTCGCGCAGCAGCAGGCTTGCATTTATAGACAGCAGCGGCTGCGCTTA 420
DB 361 CTCCTCAGTCGCGCAGCAGCAGGCTTGCATTTATAGACAGCAGCGGCTGCGCTTA 420

QY 421 TCGGAGGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAGGCTCTGACGG 480
DB 421 TCGGAGGGTCAAAAGTGTAGTGGCTGCTGGATGCTCTGTACGGGAGGCTCTGACGG 480

QY 481 ATGAGCAGTACAGGCACTGGGGCCGAGCCACCAAGCAAGATGCGGAAGCTCT 540
DB 481 ATGAGCAGTACAGGCACTGGGGCCGAGCCACCAAGCAAGATGCGGAAGCTCT 540

QY 541 TCAGTTTCACACCAAGCCTGGAACTGGACCTGCAAGGACTTGGCTCTCCAGGCCCTAAGGG 600
DB 541 TCAGTTTCACACCAAGCCTGGAACTGGACCTGCAAGGACTTGGCTCTCCAGGCCCTAAGGG 600

QY 601 AGTCCAGTCTTACTGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCCAGCAACAC 660
DB 601 AGTCCAGTCTTACTGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCCAGCAACAC 660

QY 661 TCCGCTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACACAA 720
DB 661 TCCGCTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACACAA 720

QY 721 ATACGAAAAGCAGCTTGAA 740
DB 721 ATACGAAAAGCAGCTTGAA 740

RESULT 2
ABK87966
ID ABK87966 standard; cDNA; 740 BP.
XX
AC ABK87966;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human caspase recruitment domain-5 (CARD-5) cDNA.
XX
KW Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;
KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
KW cell proliferation; gene therapy; immune disorder;
KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 54..641
FT /tag= a
FT /product= "CARD-5"
XX
PN WO200244354-A2.
XX
XX 06-JUN-2002.
PD
PF 29-NOV-2001; 2001WO-US44894.
XX
PR 01-DEC-2000; 2000US-0728721.
PR 24-APR-2001; 2001US-0841879.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2002-557538/59.
DR P-PSDB; AAU99353.
XX
XX Novel isolated murine or human caspase recruitment domain (CARD)-5
PT polypeptide, useful for treating immune disorders such as Hashimoto's
PT thyroiditis, graft rejection, allergy, glomerular nephritis,
PT tuberculosis -
XX

(NF-kappaB) transcription factor pathway and binds the CARDS of caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5 activity and NF-kappaB activation, regulate cell growth and cell death and be used in gene therapy. The CARD-5 polypeptides are useful for identifying compounds which bind to them and modulate their activity and for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides, nucleic acids, antibodies and modulators of CARD-5 expression or activity can be used to treat immune disorders such as chronic inflammatory diseases and disorders, Hashimoto's thyroiditis, graft rejection, sarcoidosis, atopic conditions (such as asthma and allergy), glomerular nephritis, human immunodeficiency virus (HIV) and bacterial infections (including tuberculosis and lepromatous leprosy) and in screening and detection assays. Modulators of CARD-5 activity or expression are also useful for treating autoimmune disorders, such as systemic lupus erythematosus and arthritis, cell depletion, neurological disorders, such as Alzheimer's disease, Parkinson's disease and spinal muscular atrophy, haematologic diseases, such as myelodysplastic syndrome and aplastic anaemia, myocardial infarction and stroke. The sequence presented is the complementary strand to the human caspase recruitment domain-5 (CARD-5) cDNA.

Sequence 740 BP; 120 A; 236 C; 238 G; 146 T; 0 other;

Query Match 100.0%; Score 740; DB 24; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGCTGACGGGGTGGCGGCGGAGCGCGCGGATCTGAGGCATGGGCG 60
 DB 740 CGCGTCCGCTGACGGGGTGGCGGCGGAGCGCGGATCTGAGGCATGGGCG 681
 QY 61 GCGCGCGCGACGCTCTGGATGGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGT 120
 DB 680 GCGCGCGCGACGCTCTGGATGGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGT 621
 QY 121 TCAAGCTGAAGCTGTGTGCGTCCGCTGCGGAGGCTTACGGCGCATCCCGCGGGCG 180
 DB 620 TCAAGCTGAAGCTGTGTGCGTCCGCTGCGGAGGCTTACGGCGCATCCCGCGGGCG 561
 QY 181 CGCTGTGTTCATGAGCGCTTGGAGCTTACCGAGAGCTGTGCTACCTTCTACCTGGAGA 240
 DB 560 CGCTGTGTTCATGAGCGCTTGGAGCTTACCGAGAGCTGTGCTACCTTCTACCTGGAGA 501
 QY 241 CTTACGGCGCGGAGCTTACCGCTTACGCTTACGCTGCGGAGGCTTACCGGAGATGGCG 300
 DB 500 CTTACGGCGCGGAGCTTACCGCTTACGCTTACGCTGCGGAGGCTTACCGGAGATGGCG 441
 QY 301 GGCAGCTGAGCGCGGCGGCGGAGGCTTGGAGCGCGGAGCTGGGATCCAGGCGCC 360
 DB 440 GGCAGCTGAGCGCGGCGGCGGAGGCTTGGAGCGCGGAGCTGGGATCCAGGCGCC 381
 QY 361 CTCCTCAGTCGGCAGCAAGCGGCTGCACTTTATAGACGACGCGGGCTGGCCTTA 420
 DB 380 CTCCTCAGTCGGCAGCAAGCGGCTGCACTTTATAGACGACGCGGGCTGGCCTTA 321
 QY 421 TCGCAGGCTTCAACAGTTGAGTGCTGTGATGCTGTACGGAGGTCCTGACGG 480
 DB 320 TCGCAGGCTTCAACAGTTGAGTGCTGTGATGCTGTACGGAGGTCCTGACGG 261
 QY 481 ATGAGCAGTACAGGCGAGTGGCGGCGGAGGCTTGGAGCGCGGAGTGGGAGGCTCT 540
 DB 260 ATGAGCAGTACAGGCGAGTGGCGGCGGAGGCTTGGAGCGCGGAGTGGGAGGCTCT 201
 QY 541 TCAGTTTTCACAGGCTGGAAGTGAAGCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGG 600
 DB 200 TCAGTTTTCACAGGCTGGAAGTGAAGCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGG 141
 QY 601 AGTCCAGTCTTACCTGGTGGAGGACTTGGAGCGGAGGCTGAGGCTCTTCCAGCAACAC 660
 DB 140 AGTCCAGTCTTACCTGGTGGAGGACTTGGAGCGGAGGCTGAGGCTCTTCCAGCAACAC 81
 QY 661 TCCGCTCAGCGCGCTGGCAATCCCAATATCATCTCTGAATCTGATCTTTTATACAAAT 720

DB 80 TCCGCTCAGCGCTGGCAATCCCAATATCATCTGAATCTGATCTTTTATACAAAT 21
 QY 721 ATACGAAAAGCCAGCTTGAA 740
 DB 20 ATACGAAAAGCCAGCTTGAA 1

RESULT 4
 ABS56032
 ID ABS56032 standard; cDNA; 740 BP.
 XX
 AC ABS56032;
 XX
 DT 10-JAN-2003 (first entry)
 XX
 DE cDNA encoding human caspase recruitment domain-5 (CARD-5).
 XX
 KW Human; caspase activity; caspase recruitment domain-5; CARD-5;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudolICE; ICEBERG; cell growth; cell death; inflammation; CARD-7;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant; gene; ss;
 KW caspase recruitment domain-7.
 XX
 KW Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 54..641
 FT /tag= a
 FT /product= "CARD-5"
 XX
 PN US2002128198-A1.
 XX
 PD 12-SEP-2002.
 XX
 XX 27-NOV-2001; 2001US-0996617.
 PF
 XX
 XX 28-JUN-1999; 99US-0340620.
 PR
 PR 27-OCT-1999; 99US-0428252.
 PR 15-AUG-2001; 2001US-0991071.
 XX
 PA (BERT/) BERTIN J.
 XX
 PI Bertin J;
 XX
 DR WPI; 2003-028967/02.
 DR P-PSDB; ABG71635.
 XX
 XX Identifying modulator of CARD-7 and CARD-5 interaction, by contacting CARD-7 and CARD-5 in presence of test compound, measuring their binding, and identifying modulator, when binding of CARD-7 to CARD-5 is altered
 PS Disclosure; Fig 7; 43pp; English.
 XX
 CC The present invention relates to methods of identifying compounds that regulate caspase activity using caspase recruitment domain-7 (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a method for identifying a compound that modulates the interaction between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta converting enzyme (pseudolICE) or ICEBERG is disclosed. CARD-7 and CARD-8 molecules are useful as modulating agents in regulating a variety of cellular processes including cell growth, cell death, and inflammation. The methods of the invention are useful for identifying compounds that have the ability to increase/decrease apoptosis, or

comprise the ability to induce caspase activation. The methods are useful for treating a disorder associated with inappropriate apoptosis or inappropriate inflammation. The methods are useful for treating disorders associated with an undesirably low rate of apoptosis such as cancer (preferably follicular lymphoma, chronic myelogenous leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral infections, autoimmune diseases caused by low levels of apoptosis (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis, and arthritis). The methods are also useful for treating disorders with undesirably high rates of apoptosis such as human immunodeficiency virus (HIV) infection, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal muscular atrophy, various forms of cerebellar degeneration, anaemia associated with chronic disease, aplastic anaemia, chronic neutropenia, myelodysplastic syndromes, myocardial infarction, stroke, and various inflammatory disorders (e.g. Crohn's disease, reactive arthritis, insulin dependent diabetes mellitus, multiple sclerosis, psoriasis, graft rejection, allergic rhinitis, food allergies, conjunctivitis, glomerular nephritis, etc). The present sequence encoding human CARD-5 is used to identify cDNA encoding CARD-7.

XX Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match 100.0%; Score 740; DB 25; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGTCCAGCGGGTCCAGCGGGGAGCGGCGGAGCGGATCTGGAGGCATGGGGC 60
 DB 1 CGCGTCCGGTCCAGCGGGTCCAGCGGGGAGCGGCGGAGCGGATCTGGAGGCATGGGGC 60

QY 61 GCGCGCGGACGCCATCTCTGGTGGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGT 120
 DB 61 GCGCGCGGACGCCATCTCTGGTGGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGT 120

QY 121 TCAAGCTGAAGTCTGTCTGGTCCGCTCGCGAGGCGCTACCGGCGATCCCGCGGGGG 180
 DB 121 TCAAGCTGAAGTCTGTCTGGTCCGCTCGCGAGGCGCTACCGGCGATCCCGCGGGGG 180

QY 181 CGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTGGAGA 240
 DB 181 CGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTGGAGA 240

QY 241 CCTACGGCGCGAGCTCACCCTGCTGCGGACATGGCGCTCGAGGAGATGGCGG 300
 DB 241 CCTACGGCGCGAGCTCACCCTGCTGCGGACATGGCGCTCGAGGAGATGGCGG 300

QY 301 GGCAGTCCAGCGCGCCACGACAGGCTCTGGAGCGCGCGAGTCCAGGCGCC 360
 DB 301 GGCAGTCCAGCGCGCCACGACAGGCTCTGGAGCGCGCGAGTCCAGGCGCC 360

QY 361 CTCCTCAGTCCGAGCAGCAGCGAGCGCTGCTTATAGCAGCAGCGGCTGGCTTA 420
 DB 361 CTCCTCAGTCCGAGCAGCAGCGAGCGCTGCTTATAGCAGCAGCGGCTGGCTTA 420

QY 421 TCGCAGGGGTCAAAACGTGTAGTGTCTGTGATCTCTGTACGGGAAGTCTTCAGCG 480
 DB 421 TCGCAGGGGTCAAAACGTGTAGTGTCTGTGATCTCTGTACGGGAAGTCTTCAGCG 480

QY 481 ATGAGCAGTACAGGCGAGTCCGGGCGAGCCCAACCAAGCAAGATGCGGAAGCTCT 540
 DB 481 ATGAGCAGTACAGGCGAGTCCGGGCGAGCCCAACCAAGCAAGATGCGGAAGCTCT 540

QY 541 TCAGTTTCACACGAGCTGMACTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGG 600
 DB 541 TCAGTTTCACACGAGCTGMACTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGG 600

QY 601 AGTCCAGTCTTACCTGGTGGAGGAGCTGGAGCGGAGCTGAGCTCTTCCAGCAAC 660
 DB 601 AGTCCAGTCTTACCTGGTGGAGGAGCTGGAGCGGAGCTGAGCTCTTCCAGCAAC 660

QY 661 TCGGTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAAT 720
 DB 661 TCGGTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAAT 720

Db 661 TCGGTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAAT 720

QY 721 ATACGAAAAGCCAGCTTGAA 740
 ||||||||||||||||||
 Db 721 ATACGAAAAGCCAGCTTGAA 740
 ||||||||||||||||||

RESULT 5
 AAD03890
 ID AAD03890 standard; cDNA; 770 BP.
 XX AAD03890;
 AC AAD03890;
 DT 02-JUL-2001 (first entry)
 DE Human target of methylation-induced silencing-1 (TMS1) cDNA.
 KW Human; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast; ss.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 75..662
 FT /*tag= a
 FT /product= "Human TMS1 protein"

XX WO200129235-A2.
 XX 26-APR-2001.
 XX 18-OCT-2000; 2000WO-US28747.
 XX 18-OCT-1999; 99US-0159975.
 XX (UYEM-) UNIV EMORY.
 XX Vertino PM;
 XX WPI; 2001-290922/30.
 XX P-PSDB; AAE00588.

Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor characterized by abnormal methylation, for treating cancer by inducing apoptosis

Claim 68; Page 113-114; 124pp; English.

The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy. TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it. The present sequence is human target of methylation-induced silencing-1 (TMS1) cDNA.

Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

Query Match 99.2%; Score 734; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 3.2e-146;
 Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGCAGCGGGGTGAGCGCGGCGGCGGAGTCTTGAGCCATGGGGCGCGCG 66

Db 28 CGGCTCAGCGGGTGAAGCGCGGAGCGCGCGGGATCTCTGAGCATGGGCGCGCGC 87
 Qy 67 GCGACGCCATCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC 126
 Db 88 GCGACGCCATCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC 147
 Qy 127 TGAAGCTGTGTGGTGGCTGGCTGGAGAGGCTTACCGCGCGATCCCGCGGGCGCGCTGC 186
 Db 148 TGAAGCTGTGTGGTGGCTGGCTGGAGAGGCTTACCGCGCGATCCCGCGGGCGCGCTGC 207
 Qy 187 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGTGAGCTTCTACCTGGAGACCTAGC 246
 Db 208 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGTGAGCTTCTACCTGGAGACCTAGC 267
 Qy 247 GCGCCGAGCTCACCGCTTAAGCTGTGCGGACATGGGCGCTGCGAGAGATGGCGGGCAGC 306
 Db 268 GCGCCGAGCTCACCGCTTAAGCTGTGCGGACATGGGCGCTGCGAGAGATGGCGGGCAGC 327
 Qy 307 TGCAGCGCGCCAGCACACGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCCCTC 366
 Db 328 TGCAGCGCGCCAGCACACGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCCCTC 387
 Qy 367 AGTGGCAGCCAAAGCCAGGCTGCACTTTATAGACAGCACCGGGCTGGCGCTTATGCGGA 426
 Db 388 AGTGGCAGCCAAAGCCAGGCTGCACTTTATAGACAGCACCGGGCTGGCGCTTATGCGGA 447
 Qy 427 GGGTCACAAAGTGTGAGTGGCTGTGAGTGTCTGTACGGGAAGTCTCTGACGAGATGAGC 486
 Db 448 GGGTCACAAAGTGTGAGTGGCTGTGAGTGTCTGTACGGGAAGTCTCTGACGAGATGAGC 507
 Qy 487 AGTACAGCAGTGTGGGCGCGCGCCAGCACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
 Db 508 AGTACAGCAGTGTGGGCGCGCGCCAGCACCAAGCAAGATGCGGAAGCTCTTCAGTT 567
 Qy 547 TCACACAGCCTGGAACTGGACCTGCAAGGACTGTCTCTCCAGGCGCTTAAGGAGTCCC 606
 Db 568 TCACACAGCCTGGAACTGGACCTGCAAGGACTGTCTCTCCAGGCGCTTAAGGAGTCCC 627
 Qy 607 AGTCTACCTGTGGAGGACTGGAGCGAGCTGAGGCTCTCTCCAGCAACACTCCCGT 666
 Db 628 AGTCTACCTGTGGAGGACTGGAGCGAGCTGAGGCTCTCTCCAGCAACACTCCCGT 687
 Qy 667 CAGCCCTCGCAATCCCAATATCTGAACTGATCTGATCTTTTATACAAATATACGA 726
 Db 688 CAGCCCTCGCAATCCCAATATCTGAACTGATCTGATCTTTTATACAAATATACGA 747
 Qy 727 AAAGCCAGCTTGAA 740
 Db 748 AAAGCCAGCTTGAA 761

RESULT 6

AA47126
 ID AAL47126 standard; DNA; 782 BP.
 XX
 AC AAL47126;
 XX
 DT 20-AUG-2002 (first entry)
 XX
 DE Pyrin domain containing protein Pycard coding sequence.
 XX
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antiporiatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteoprotic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis; gene; ds.
 OS Unidentified.
 XX
 XX WO200240668-A2.
 PN

XX 23-MAY-2002.
 XX 30-OCT-2001; 2001WO-EP12545.
 PF 15-NOV-2000; 2000DE-1056687.
 PR 30-NOV-2000; 2000DE-1059595.
 XX (APOT-) APOTECH RES & DEV LTD.
 PA Tschoep J, Martinon F;
 XX WPI; 2002-427093/45.
 DR P-PSDB; AAO17854.
 XX New DNA encoding protein with pyrin domain, useful for treating
 PT diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies -
 XX Claim 5; Fig 1; 116pp; German.
 PS The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a coding sequence of
 CC the invention.
 XX
 SQ Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other;

Query Match 99.2%; Score 734; DB 24; Length 782;
 Best Local Similarity 100.0%; Pred. No. 3.2e-146;
 Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 CGGCTGACAGCGGGTGAAGCGCGGAGCTCTGGAGCATGGGCGCGCGC 66
 Db 40 CGGCTGACAGCGGGTGAAGCGCGGAGCTCTGGAGCATGGGCGCGCGC 99
 Qy 67 GCGACGCCATCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC 126
 Db 100 GCGACGCCATCTGGATGGCTGGAGAACTGACCGCGAGGAGCTCAAGAGTTCAAGC 159
 Qy 127 TGAAGCTGTGTGGTGGCTGGCTGGAGAGGCTTACCGCGCGATCCCGCGGGCGCGCTGC 186
 Db 160 TGAAGCTGTGTGGTGGCTGGCTGGAGAGGCTTACCGCGCGATCCCGCGGGCGCGCTGC 219
 Qy 187 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGTGAGCTTCTACCTGGAGACCTAGC 246
 Db 220 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGTGAGCTTCTACCTGGAGACCTAGC 279
 Qy 247 GCGCCGAGCTCACCGCTTAAGCTGTGCGGACATGGGCTTGCAGAGATGGCGGGCAGC 306
 Db 280 GCGCCGAGCTCACCGCTTAAGCTGTGCGGACATGGGCTTGCAGAGATGGCGGGCAGC 339
 Qy 307 TGCAGCGCGCCAGCACACGAGGCTTGCAGCGCGCGAGCTGGGATCCAGGCGCCCTCCTC 366
 Db 340 TGCAGCGCGCCAGCACACGAGGCTTGCAGCGCGCGAGCTGGGATCCAGGCGCCCTCCTC 399
 Qy 367 AGTGGCAGCCAAAGCCAGGCTGCACTTTATAGACAGCACCGGGCTGGCTTATGCGGA 426
 Db 400 AGTGGCAGCCAAAGCCAGGCTGCACTTTATAGACAGCACCGGGCTGGCTTATGCGGA 459
 Qy 427 GGGTCACAAAGTGTGAGTGGCTGTGAGTGTCTGTACGGGAAGTCTCTGACGAGATGAGC 486
 Db 460 GGGTCACAAAGTGTGAGTGGCTGTGAGTGTCTGTACGGGAAGTCTCTGACGAGATGAGC 519
 Qy 487 AGTACAGCAGTGTGGGCGCGCGCCAGCACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
 Db 520 AGTACAGCAGTGTGGGCGCGCGCCAGCACCAAGCAAGATGCGGAAGCTCTTCAGTT 579

QY 547 TCACACACGCTGGAACTGGAGCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 606
 Db 580 TCACACACGCTGGAACTGGAGCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 639
 QY 607 AGTCCTACCTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
 Db 640 AGTCCTACCTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCGCT 699
 QY 667 CAGCCCTCGCAATCCACCAATCATCTGAATCTGTATCTTTTATACACAATATACGA 726
 Db 700 CAGCCCTCGCAATCCACCAATCATCTGAATCTGTATCTTTTATACACAATATACGA 759
 QY 727 AAAGCAGCTTGA 740
 Db 760 AAAGCAGCTTGA 773

RESULT 7

AAS59817
 ID AAS59817 standard; cDNA; 745 BP.

AC AAS59817;

DT 16-JAN-2002 (first entry)

DE Human novel cytokine encoding cDNA 790CIP2B_1 #1.

XX Human; ss; cytokine; cell proliferation; cell differentiation;
 KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement.

XX Homo sapiens.

OS WO200175093-A1.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US10484.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PR 22-SEP-2000; 2000US-0668680.

PR 23-OCT-2000; 2000US-0695618.

PR 30-NOV-2000; 2000US-0728711.

PR 14-MAR-2001; 2000US-0728711.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;

PI Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;

XX WPI; 2001-626432/72.

DR P-PSDB; AAU68525.

XX New polypeptides and nucleic acids, useful for diagnosis, treatment of
 PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
 PT degenerative disorders, cancer and promoting wound healing -
 XX Claim 1; Page 150-151; 336pp; English.

XX The invention relates to isolated human polypeptides (which may be
 CC cytokines) and the polynucleotides encoding them. The protein is useful
 CC for identifying a compound which binds to it (e.g. modulators, agonists

CC and antagonists). The polynucleotides are useful as an array for mismatch
 CC detection. The proteins and nucleic acids are useful as nutritional
 CC sources or supplements. The protein exhibits exhibits activity relating
 CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity, immune stimulating or immune
 CC suppressing and activin or inhibin related activities. The proteins (and
 CC antibodies raised against them) and nucleic acids are therefore useful in
 CC the diagnosis and treatment of diseases and disorders such as cancer,
 CC central and peripheral nervous system diseases and neuropathies,
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
 CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
 CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth, and in tissue repair, healing of burns, incisions, ulcers, for
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
 CC periodontal disease, lung or liver fibrosis, reperfusion injury in
 CC various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
 CC such as asthma or other respiratory problems, coagulation disorders,
 CC haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory
 CC bowel disease, viral infection and are useful in altering bodily
 CC characteristics. The present sequence encodes a novel protein of the
 CC invention.

XX Sequence 745 BP; 157 A; 234 C; 235 G; 119 T; 0 other;

QY Query Match 99.1%; Score 733; DB 22; Length 745;

Db Best Local Similarity 100.0%; Pred. No. 5.2e-146;

QY Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GGCTCAGCGGGTGCAGCGCGGCGGATCTCTGAGCATCTGGGCGCGCGCGG 67

Db 2 GGCTCAGCGGGTGCAGCGCGGCGGATCTCTGAGCATCTGGGCGCGCGCGG 61

QY 68 CGACGCCATCTCTGGATGGCTGGAGAACTGACCCCGGAGAGCTCAAGAACTTCAAGCT 127

Db 62 CGACGCCATCTCTGGATGGCTGGAGAACTGACCCCGGAGAGCTCAAGAACTTCAAGCT 121

QY 128 GAAGCTGTCTGCTGCGGCGGAGGCTACGGGCGCATCCCGCGGGCGCGCTGT 187

Db 122 GAAGCTGTCTGCTGCGGCGGAGGCTACGGGCGCATCCCGCGGGCGCGCTGT 181

QY 188 GTCCATGGAGCGCTTGGACCTCACGACAAGCTGTACGTTCTACCTGGAGACCTACCG 247

Db 182 GTCCATGGAGCGCTTGGACCTCACGACAAGCTGTACGTTCTACCTGGAGACCTACCG 241

QY 248 CGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGCAGGAGATGCCCGGCGAGCT 307

Db 242 CGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGCAGGAGATGCCCGGCGAGCT 301

QY 308 GCAGCGGCGCACGACACGAGGCTCTGGAGCGCGCGGAGCTGGGATCCAGGCCCTCTCTCA 367

Db 302 GCAGCGGCGCACGACACGAGGCTCTGGAGCGCGCGGAGCTGGGATCCAGGCCCTCTCTCA 361

QY 368 GTCCGAGCGCACGAGCGCTGACCTTTATAGACGACACCGGCTGCGCTTATCGCGAG 427

Db 362 GTCCGAGCGCACGAGCGCTGACCTTTATAGACGACACCGGCTGCGCTTATCGCGAG 421

QY 428 GGTCAACAAAGTTGAGTGGCTGCTGTGATCGGAGAGGCTCTCACCGATGAGCA 487

Db 422 GGTCAACAAAGTTGAGTGGCTGCTGTGATCGGAGAGGCTCTCACCGATGAGCA 481

QY 488 GTACGAGCGAGTGGCGGCGGAGCCACCAACCAAGCAAGATCGGAAAGCTTTCAGTTT 547

Db 482 GTACGAGCGAGTGGCGGCGGAGCCACCAACCAAGCAAGATCGGAAAGCTTTCAGTTT 541

QY 548 CACACGAGCTGGAACCTGGAACCTGGAACCTGCTCTCTCCAGGCCCTTAAGGAGTCCCA 607

Db 542 CACACGAGCTGGAACCTGGAACCTGGAACCTGCTCTCTCCAGGCCCTTAAGGAGTCCCA 601

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PS Claim 1; Page 2955; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 806 BP; 163 A; 253 C; 260 G; 128 T; 2 other;
Query Match 97.5%; Score 721.2; DB 22; Length 806;
Best Local Similarity 99.6%; Pred. No. 1.7e-143;
Matches 731; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
7 CGGCTGCAGCGGGGTGAGCGCGGCGGAGCTCTGGAGCATGGCGCGCGCG 66
68 CGGCTGCAGCGGGGTGAGCGCGGCGGAGCTCTGGAGCATGGCGCGCGCG 127
67 GCGAGCGCATCTGATGCGGTGGAGAACCTGACCGCGCGGAGGCTCAAGAGTTCAAGC 126
128 GCGAGCGCATCTGATGCGGTGGAGAACCTGACCGCGCGGAGGCTCAAGAGTTCAAGC 187
127 TGAAGTGTGTCGGTGGCGGCGGAGGCTACGGCGCATCCGCGGGCGCGCTGC 186
188 TGAAGTGTGTCGGTGGCGGCGGAGGCTACGGCGCATCCGCGGGCGCGCTGC 247
187 TGTCCATGACGCTTGGACCTCACGCAAGCTGTCTACCTTCTACCTGGAGACTACG 246
248 TGTCCATGACGCTTGGACCTCACGCAAGCTGTCTACCTTCTACCTGGAGACTACG 307
247 GCGCGGAGTCTACCGTAAACGTCGTGCGGACATGGGCTTGAGGAGATGGCGGCGAGC 306
308 GCGCGGAGTCTACCGTAAACGTCGTGCGGACATGGGCTTGAGGAGATGGCGGCGAGC 367
307 TGCAGGCGGCGGACGACACGAGGCTCTGGAGCGCGGCTGAGTCCAGGCGGCTCTC 366
368 TGCAGGCGGCGGACGACACGAGGCTCTGGAGCGCGGCTGAGTCCAGGCGGCTCTC 427
367 AGTTCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 426
428 AGTTCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 487
427 GGGTCAAAACGTTGAGTGGTGTCTGTGATGCTGTCTGTACGGGAGGTCCTGAGGATGAGC 486
488 GGGTCAAAACGTTGAGTGGTGTCTGTGATGCTGTCTGTACGGGAGGTCCTGAGGATGAGC 547
487 AGTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 546
548 AGTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 607
547 TCACACGAGCTGGAAGTGAAGCTGCAAGCTTGTCTCTCCAGGCGGCTTAAGGAGTCC 606
608 TCACACGAGCTGGAAGTGAAGCTGCAAGCTTGTCTCTCCAGGCGGCTTAAGGAGTCC 667
607 AGTCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666

Db 668 AGTCTTACCTGGTGGAGGACCTKGAGCGAGCTGAGGCTCTTCCAGCAACACTCCGCT 727
Qy 667 CAGCCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
Db 728 CA-CCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATACGA 786
Qy 727 AAAGCCAGCTTGAA 740
Db 787 AAAGCCAGCTTGAA 800
RESULT 10
AAC77884
ID AAC77884 standard; cDNA; 811 BP.
XX
AC AAC77884;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:278.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05882.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587533/55.
DR P-PSDB; AAB43675.
PT Novel isolated nucleic acids comprising sequences encoding peptides
PS useful for treating or diagnosing e.g. cancer -
XX
PS Claim 1; Page 841; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerrary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to

CC	AACT8457 and AAB44240 represent sequences used in the exemplification of the present invention.
XX	
SQ	Sequence 811 BP; 170 A; 253 C; 259 G; 128 T; 1 other;
	Query Match 94.3%; Score 697.6; DB 21; Length 811;
	Best Local Similarity 99.5%; Pred. No. 1.6e-138;
	Matches 730; Conservative 1; Mismatches 0; Indels 3; Gaps 3;
Qy	7 CGGCTGACGGGGTGAGCGGGGAGCGCGGGATCCTGAGGCATGGGGCGCGCGC 66
Db	68 CGGCTGACGGGGTGAGCGGGGAGCGCGGGATCCTGAGGCATGGGGCGCGCGC 127
Qy	67 GCGACGCCATCTGGATGGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 126
Db	128 GCGACGCCATCTGGATGGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 187
Qy	127 TGAAGCTGTGTCGGTCCGCTGCGCGAGGGCTACGCGCGCATCCCGGGGCGCGCTGC 186
Db	188 TGAAGCTGTGTCGGTCCGCTGCGCGAGGGCTACGCGCGCATCCCGGGGCGCGCTGC 247
Qy	187 TGTCCATGACCGCTTGGACCTCACGCAAGCTGGTCACTTCTACCTGGAGAGCTACG 246
Db	248 TGTCCATGACCGCTTGGACCTCACGCAAGCTGGTCACTTCTACCTGGAGAGCTACG 307
Qy	247 GCGCCGAGCTCACCGCTAAACGTGCTGCGCGACATGGGCTCGAGGAGATGGCGGGCAGC 306
Db	308 GCGCCGAGCTCACCGCTAAACGTGCTGCGCGACATGGGCTCGAGGAGATGGCGGGCAGC 367
Qy	307 TGCAGCGGCCACGACACAGGGCTGTGGAGCGCGCAGCTGGGATCCAGGCCCTCTCTC 366
Db	368 TGCAGCGGCCACGACACAGGGCTGTGGAGCGCGCAGCTGGGATCCAGGCCCTCTCTC 427
Qy	367 AGTCGCGCAGCAGCAGCGCTGCACTTTATAGACGACGCGGCTGCGCTTATCGGA 426
Db	428 AGTCGCGCAGCAGCAGCGCTGCACTTTATAGACGACGCGGCTGCGCTTATCGGA 487
Qy	427 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTACGGGAAGGTCTCGAGGATGAGC 486
Db	488 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTACGGGAAGGTCTCGAGGATGAGC 547
Qy	487 AGTACAGCAGTGGCGGCGGAGCCCAACACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
Db	548 AGTACAGCAGTGGC-GGCGGAGCCCAACACCAAGCAAGATGCGGAAGCTCTTCAGTT 606
Qy	547 TCACACCGCTGGAACCTGCAAGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCC 606
Db	607 TCACACCGCTGGAACCTGCAAGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCC 666
Qy	607 AGTCCTACCTGGTGGAGGACCTGGAGCGAGCTGAGGCTCCTTCCAGCAACACTCCGGT 666
Db	667 AGTCCTACCTGGTGGAGGACCTGGAGC-GAGCTGAGGCTCCTTCCAGCAACACTCCGGT 725
Qy	667 CAGCCCTCGCAATCCCAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
Db	726 CA-SCCCTGGCAATCCCAATCATCTGAATCTGATCTTTTATACAAATATACGA 784
Qy	727 AAAGCCAGCTTGA 740
Db	785 AAAGCCAGCTTGA 798
RESULT 11	
AAD03906	
ID	AAD03906 standard; cDNA; 713 BP.
XX	
AC	AAD03906;
XX	
DT	02-JUL-2001 (first entry)
XX	
DE	Alternatively spliced form of human TMS1 cDNA (lacking exon2).
XX	
KW	Human; target of methylation-induced silencing-1; TMS1; cytostatic;

KW	antiproliferative; apoptosis inducer; gene therapy; CpG island; caspase-recruiting domain; CARD; cancer; breast; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	75..605
FT	/tag= a
FT	/product= "Alternatively spliced form of human TMS1 protein lacking exon2"
XX	
PN	WO200129235-A2.
XX	
PD	26-APR-2001.
XX	
PF	18-OCT-2000; 2000WO-US28747.
XX	
PR	18-OCT-1999; 99US-0159975.
XX	
PA	(UYEM-) UNIV EMORY.
XX	
PI	Vertino PM;
XX	
DR	WPI; 2001-290922/30.
DR	P-PSDB; AAE00594.
XX	
PT	Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor characterized by abnormal methylation, for treating cancer by inducing apoptosis -
XX	
PS	Claim 68; Page 122; 124pp; English.
XX	
CC	The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC	TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it.
CC	The present sequence is alternatively spliced form of human target of methylation-induced silencing-1 (TMS1) cDNA lacking exon2.
XX	
SQ	Sequence 713 BP; 148 A; 224 C; 223 G; 118 T; 0 other;
	Query Match 82.4%; Score 610; DB 22; Length 713;
	Best Local Similarity 92.2%; Pred. No. 5.6e-120;
	Matches 677; Conservative 0; Mismatches 0; Indels 57; Gaps 1;
Qy	7 CGGCTGACGGGGTGAGCGGGGAGCGCGGGATCCTGGAGCCATGGGGCGCGCGC 66
Db	28 CGGCTGACGGGGTGAGCGGGGAGCGCGGGATCCTGGAGCCATGGGGCGCGCGC 87
Qy	67 GCGAGCCCATCTGGATGCGTGGAGAACTTGACCGCGGAGAGCTCAAGAAAGTTCAAGC 126
Db	88 GCGAGCCCATCTGGATGCGTGGAGAACTTGACCGCGGAGAGCTCAAGAAAGTTCAAGC 147
Qy	127 TGAAGCTGTGTCGGTCCGCTGCGGAGGGCTACGGCGGCATCCCGGGGCGCGCTGC 186
Db	148 TGAAGCTGTGTCGGTCCGCTGCGGAGGGCTACGGCGGCATCCCGGGGCGCGCTGC 207
Qy	187 TGTCCATGAGCCCTTGGACCTCACCGACAAAGCTGGTCACTTCTACCTGGAGAGCTACG 246
Db	208 TGTCCATGAGCCCTTGGACCTCACCGACAAAGCTGGTCACTTCTACCTGGAGAGCTACG 267
Qy	247 GCGCGGAGCTCACCGCTTAACGTGTGCGCGCATGGGCTCGCAGGAGATGGCCGGCGAGC 306

Db 268 GCGCCGAGCTACCGCTAACGTGCTGCGCGACATGGGCCCTGCAGAGATGGCCGGCGAGC 327
QY 307 TGCAGCGCGCCACGACACAGGCGCTCTGGAGCGCGCGCAGCTGGATCCAGGCGCCCTCTCTC 366
Db 328 TGCAGCGCGCCACGACACG-----GCGCTGCACTTTATAGACCGACCGGGCTGGCTTATCGCGA 347
QY 367 AGTCGGCAGCCAAAGCCAGCCCTGCACTTTATAGACCGACCGGGCTGGCTTATCGCGA 426
Db 348 -----GCGCTGCACTTTATAGACCGACCGGGCTGGCTTATCGCGA 390
QY 427 GGGTCACAAACCTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGC 486
Db 391 GGGTCACAAACCTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGC 450
QY 487 AGTACGAGCAGCTGGCGGCGCGACCCACCAACCCAAAGCAAGATGGCGAAGCTCTTCAGTT 546
Db 451 AGTACGAGCAGCTGGCGGCGCGACCCACCAACCCAAAGCAAGATGGCGAAGCTCTTCAGTT 510
QY 547 TCACACGAGCCTGGAACCTGGAACCTGGAAGCAAGCTTGTCTCCAGGCGCTTAAGGGAGTCCC 606
Db 511 TCACACGAGCCTGGAACCTGGAAGCAAGCTTGTCTCCAGGCGCTTAAGGGAGTCCC 570
QY 607 AGTCTACTCTGCTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGT 666
Db 571 AGTCTACTCTGCTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGT 630
QY 667 CAGCCCTGGCAATCCACCAATCATCTGATCTGATCTTTTATACAAATATACGA 726
Db 631 CAGCCCTGGCAATCCACCAATCATCTGATCTTTTATACAAATATACGA 690
QY 727 AAAGCCAGCTTGAA 740
Db 691 AAAGCCAGCTTGAA 704

RESULT 12

AAH34053
ID AAH34053 standard; cDNA; 639 BP.
XX AAH34053;
AC AAH34053;
XX 03-SEP-2001 (first entry)
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1135.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX Homo sapiens.
OS
XX WO200122920-A2.
PN
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX WPI; 2001-235357/24.
DR P-PSDB; AAG74648.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 2955; 9803pp; English.
PS
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
CC

CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 639 BP; 146 A; 198 C; 189 G; 106 T; 0 other;

QY Query Match 71.8%; Score 531; DB 22; Length 639;
Db Best Local Similarity 94.9%; Pred. No. 2.9e-103;
QY Matches 595; Conservative 0; Mismatches 20; Indels 12; Gaps 4;

QY 124 AGCTGAAGTGTCTGCTGGTGGCTGCGGAGGGCTACGGGGCGATCCCGCG-----G 176
Db 2 AGCTGAAGTGTCTGCTGGTGGCTGCGGAGGGCTACGGGGCGCGACGCCATCTCTG 61

QY 177 GCGCGCTGCTGTCTCATGTGAGCGCTTGACCTTGCACCGA---CAAGCTGTGAGTCTAC 233
Db 62 GATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGTGTGAGTCTCTAC 121

QY 234 CTGGAGACCTACGGCGCGAGCTCACCGCTAACGTGCTGCGGACATGGCGCTGAGGAG 293
Db 122 CTGGAGACCTACGGCGCGAGCTCACCGCTAACGTGCTGCGGACATGGCGCTGAGGAG 181

QY 294 ATGGCGCGGAGCTGACGAGCGGCGCACGACACAGGGCTCTGGAGCGCGCGAGTGGGATC 353
Db 182 ATGGCGCGGAGCTGACGAGCGGCGCACGACACAGGGCTCTGGAGCGCGCGAGTGGGATC 241

QY 354 CAGGCGCGCTCTCAGTGGCGGAGCGCAAGCGCTGCACTTTATAGACCGACCGGGCT 413
Db 242 CAGGCGCGCTCTCAGTGGCGGAGCGCAAGCGCTGCACTTTATAGACCGACCGGGCT 301

QY 414 GCGCTTATCGGAGGGTCAACAAACGTTGAGTGGCTGCTGAGTCTGTACGGGAGGTC 473
Db 302 GCGCTTATCGGAGGGTCAACAAACGTTGAGTGGCTGCTGAGTCTGTACGGGAGGTC 361

QY 474 CTGACGGATGAGCAGTACGAGCGGCGCGAGCGCGCGAGCCACCAACCAAGCAAGATGCGG 533
Db 362 CTGACGGATGAGCAGTACGAGCGGCGCGAGCGCGAGCGCGAGCGCGAGTGGCGG 420

QY 534 AAGCTCTTCAGTTTACACACGAGCTTGGAACTGGAACTGGAACTGGCTTCCAGGCG 593
Db 421 AAGCTCTTCAGTTTACACACGAGCTTGGAACTGGAACTGGAACTGGCTTCCAGGCG 480

QY 594 CTAAGGGAGTCCAGTCTCTACCTGTTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCCA 653
Db 481 CTAAGGGAGTCCAGTCTCTACCTGTTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCCA 540

QY 654 GCAACACTCCGGTACGCGCGCTGGCAATCCCAAAATCATCTGATCTGATCTTTTAT 713
Db 541 GCAACACTCCGGTACGCGCGCTGGCAATCCCAAAATCATCTGATCTGATCTTTTAT 599

QY 714 ACACATATACGAAAGCGAGCTTGAA 740
Db 600 ACACATATACGAAAGCGAGCTTGAA 626

RESULT 13

AAH98638
ID AAC98638 standard; cDNA; 432 BP.
XX

CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is alternatively spliced form of human target of
CC methylation-induced silencing-1 (TMS1) cDNA lacking exon3.
CC Note: The present sequence designated as SEQ ID NO:26, is referred
CC as DNA throughout the specification, however this sequence has been
CC referred as protein in claim 106.
XX
SQ Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other;

Query Match 51.1%; Score 378; DB 22; Length 405;
Best Local Similarity 100.0%; Pred. No. 6.6e-71;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CGGCTGACGGGGTGGAGCGCGGAGCGCGGGGATCTGGAGCCATGGGGCGCGGC 66
Db 28 CGGCTGACGGGGTGGAGCGCGGAGCGCGGGGATCTGGAGCCATGGGGCGCGGC 87
QY 67 GCGAGCCATCTGGATGCTGGAGCACTGACCGCGGAGGCTCAAGAGTTCAAGC 126
Db 88 GCGAGCCATCTGGATGCTGGAGCACTGACCGCGGAGGCTCAAGAGTTCAAGC 147
QY 127 TGAAGCTGCTGCGGTCGCGCGAGGGCTACGGGCGCATCCCGGGCGCGCTGC 186
Db 148 TGAAGCTGCTGCGGTCGCGCGAGGGCTACGGGCGCATCCCGGGCGCGCTGC 207
QY 187 TGTCCATGACCGCTTGGACCTCACCGACAAGTGTGCTGAGTTCTACCTGGAGACTAGC 246
Db 208 TGTCCATGACCGCTTGGACCTCACCGACAAGTGTGCTGAGTTCTACCTGGAGACTAGC 267
QY 247 GCGCGAGCTACCGCTACAGTCTGCGCGACATGGGCGCTGAGGATGGCGGGCAGC 306
Db 268 GCGCGAGCTACCGCTACAGTCTGCGCGACATGGGCGCTGAGGATGGCGGGCAGC 327
QY 307 TCGAGCGCGCCAGCACACCGGCTCTGGAGCGCGCGAGCTGGATCCAGGCGCCCTCCTC 366
Db 328 TCGAGCGCGCCAGCACACCGGCTCTGGAGCGCGCGAGCTGGATCCAGGCGCCCTCCTC 387
QY 367 AGTCGCGACCCAGCCAG 384
Db 388 AGTCGCGACCCAGCCAG 405

RESULT 15
AAD03889
ID AAD03889 standard; DNA; 2821 BP.
XX
AC AAD03889;
XX DT 02-JUL-2001 (first entry)
XX DE Human target of methylation-induced silencing-1 (TMS1) genomic DNA.
XX KW Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_feature 1100..1725
FT /tag= a
FT /label= CpG island
FT /note= "This region is specifically claimed in claim 71"
FT 1256..2361
FT /tag= b
FT CDS
FT /product= "Human TMS1 protein"
FT /note= "CDS is interrupted by 2 introns"
FT 1256..1529
FT exon

FT intron /tag= c
FT /number= 1
FT 1530..1742
FT /tag= d
FT /number= 1
FT 1743..1799
FT /tag= e
FT /number= 2
FT 1800..2104
FT /tag= f
FT /number= 2
FT 2105..2361
FT /tag= g
FT /number= 3
XX WO200129235-A2.
XX 26-APR-2001.
XX 18-OCT-2000; 2000WO-US28747.
XX 18-OCT-1999; 99US-0159975.
XX (UYEM-) UNIV EMORY.
XX Vertino PM;
XX WPI; 2001-290922/30.
XX P-PSDB; AAE00588.
XX Novel gene TMS1, transcriptionally silenced due to increased
XX methylation useful for identifying subject at risk of developing tumor
XX characterized by abnormal methylation, for treating cancer by inducing
XX apoptosis
XX Claim 68; Page 112-113; 124pp; English.
XX The invention relates to identification of target of methylation-induced
XX silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
XX abnormal methylation of a CpG island in its 5' regulatory region. TMS1
XX consists of a carboxy terminal caspase-recruiting domain (CARD) and
XX plays a role in induction of apoptosis. TMS1 gene and protein are useful
XX as tools for diagnosing and treating a subject at risk of developing
XX cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
XX abnormally low levels of TMS1 expression products. Unique fragments of
XX TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
XX TMS1 molecule is also useful for treating abnormal cell proliferation by
XX increasing TMS1 polypeptide level to an above normal level. The CpG
XX island region of TMS1 or its fragments are used to study the methylation
XX patterns apart from any coding region contained in it.
XX The present sequence is human target of methylation-induced silencing-1
XX (TMS1) genomic DNA.
XX SQ Sequence 2821 BP; 614 A; 864 C; 757 G; 586 T; 0 other;
Query Match 48.6%; Score 360; DB 22; Length 2821;
Best Local Similarity 100.0%; Pred. No. 5.6e-67;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 381 CCAGGCGCTGCATTTATAGACACGACCGGGTGGCTTATCCGAGGTTCAAAACGTT 440
Db 2101 CCAGGCGCTGCATTTATAGACACGACCGGGTGGCTTATCCGAGGTTCAAAACGTT 2160
QY 441 GAGTGGCTGCTGGATGCTCTGTACGGAGGTCCTGACGGATGAGCAGTACAGGCGAGTG 500
Db 2161 GAGTGGCTGCTGGATGCTCTGTACGGAGGTCCTGACGGATGAGCAGTACAGGCGAGTG 2220
QY 501 CGGGCGGACCCACCAACCAAGCAAGATGCCGAAGCTCTTCAGTTTCAACAGCCTGG 560
Db 2221 CGGGCGGACCCACCAACCAAGCAAGATGCCGAAGCTCTTCAGTTTCAACAGCCTGG 2280
QY 561 AACTGAGCTCGAAGGACTTGTCTCCAGGCCCTTAAGGAGTCCAGTCTTACTACCTGGTG 620
FT

Db	2281	AACTGGACCTGCAAGGACTTGCTCTCCAGGCCCTAAGGGAGTCCCAAGTCCTACCTGGTG	2340
Qy	621	GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGTCAGCCCTGGCAAT	680
Db	2341	GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGTCAGCCCTGGCAAT	2400
Qy	681	CCCACCAATCATCCTGAATCTGATCTTTTATACAAATATACGAAAAGCCAGCTTGAA	740
Db	2401	CCCACCAATCATCCTGAATCTGATCTTTTATACAAATATACGAAAAGCCAGCTTGAA	2460

Search completed: January 29, 2004, 03:29:23
Job time : 184.487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:01:52 ; Search time 1474.73 Seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-7

Perfect score: 740

Sequence: 1 cgcgtccgctgcagcgggg.....atcgaaagccagcttga 740

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729.4	98.6	930	12	BM459678 AGENCOURT
2	727	98.2	804	12	BM926438 AGENCOURT
3	722	97.6	767	12	BM910918 AGENCOURT
4	721.2	97.5	792	12	BM549665 AGENCOURT

5	713.4	96.4	750	12	EG764161
6	704.8	95.2	725	14	CA309883
7	690.8	93.4	856	12	EG767422
8	688.6	93.1	748	10	EG337806
9	670.4	90.6	678	9	A1148558
10	667.6	90.2	741	12	EM051141
11	660	89.2	676	12	BM972785
12	654.2	88.4	675	10	BE908204
13	647.4	87.5	744	10	EG255521
14	639	86.4	687	10	BE909218
15	617.2	83.4	638	12	BM998354
16	610	82.4	1079	12	BM456838
17	606	81.9	633	12	EM681310
18	604.6	81.7	1201	13	EX356773
19	581.2	78.5	830	10	BF238010
20	553.8	74.8	621	12	BI468249
21	546	73.8	866	10	EG684386
22	544.6	73.6	634	9	A1991236
23	539	72.8	878	13	BU501178
24	536	72.4	888	10	BE906296
25	533	72.0	615	9	A1346818
26	511.4	69.1	542	12	BQ045282
27	488.4	66.0	518	13	BU626120
28	477.4	64.5	567	9	AA528254
29	471	63.6	572	14	CA439478
30	469.2	63.4	520	12	BI468250
31	468.4	63.3	486	12	BQ012745
32	463	62.6	500	10	AW973297
33	450.4	60.9	481	9	AA573948
34	446	60.3	446	12	BM726665
35	441.8	59.7	465	9	AA582937
36	431.6	58.3	700	14	CB429243
37	430.4	58.2	698	14	CB452677
38	426	57.6	695	14	CB422326
39	420.6	56.8	1093	12	BM920038
40	409.8	55.4	421	9	AA278825
41	408	55.1	671	14	CB421971
42	407	55.0	669	14	CB429601
43	390	52.7	565	12	BM362328
44	385.4	52.1	583	14	CB436146
45	382	51.6	395	9	AI262374

ALIGNMENTS

RESULT 1
BM459678
LOCUS
DEFINITION AGENCOURT_6417789 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534308
5', mRNA sequence.
ACCESSION BM459678.1 GI:18508718
VERSION BM459678
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 930)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12220 row: j column: 05
High quality sequence stop: 670.

FEATURES	Location/Qualifiers	
	1. .930	
source	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5534308"	
	/issue_type="leiomyosarcoma"	
	/lab_host="DH10B (phage-resistant)"	
	/clone_lib="NIH_MGC_71"	
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. "	
BASE COUNT		182 a 310 c 263 g 172 t 3 others
ORIGIN		
Query Match		98.6%; Score 729.4; DB 12; Length 930;
Best Local Similarity		99.7%; Pred. No. 8.5e-154;
Matches		730; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	9	GCTGCAGCGGGTGAGCGCGGAGCGCGGGGATCTGGAGCCATGGCGCGCGCGC 68
Db	1	GCTGCAGCGGGTGAGCGCGGAGCGCGGGGATCTGGAGCCATGGCGCGCGCGC 60
QY	69	GACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAGTTTCAAGCTG 128
Db	61	GACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAGTTTCAAGCTG 120
QY	129	AGCTGCTGTCGCTGCGGTGCGGAGGGCTACGGCGGATCCCGCGGGCGGCGTGTG 188
Db	121	AGCTGCTGTCGCTGCGGTGCGGAGGGCTACGGCGGATCCCGCGGGCGGCGTGTG 180
QY	189	TCCATGAGCGCTTGACCTCACCGACAAGTGTGCTTACCTGAGACCTACGCG 248
Db	181	TCCATGAGCGCTTGACCTCACCGACAAGTGTGCTTACCTGAGACCTACGCG 240
QY	249	GCCGAGCTCACCGCTAACGCTGTGCGGACATGGGCTTCAGAGAGATGCCGGGAGCTG 308
Db	241	GCCGAGCTCACCGCTAACGCTGTGCGGACATGGGCTTCAGAGAGATGCCGGGAGCTG 300
QY	309	CAGCGGCGCACGACAGGGCTCTGAGCGCGGCGAGCTGGGATCAGGCCCTCTCTCAG 368
Db	301	CAGCGGCGCACGACAGGGCTCTGAGCGCGGCGAGCTGGGATCAGGCCCTCTCTCAG 360
QY	369	TCGGCAGCAAGCCAGCGCTGCACCTTTATAGACCAAGCGGCTGCGCTTATCGGAGG 428
Db	361	TCGGCAGCAAGCCAGCGCTGCACCTTTATAGACCAAGCGGCTGCGCTTATCGGAGG 420
QY	429	GTCACAAAGCTTGAGTGGCTGTGATGCTGTGTACGGGAAGTCTCTGACGATGAGCAG 488
Db	421	GTCACAAAGCTTGAGTGGCTGTGATGCTGTGTACGGGAAGTCTCTGACGATGAGCAG 480
QY	489	TACCAAGAGTGGGGCGAGCCACCAACCCAGCAAGATGGGAAGCTCTTCAAGTTTC 548
Db	481	TACCAAGAGTGGGGCGAGCCACCAACCCAGCAAGATGGGAAGCTCTTCAAGTTTC 540
QY	549	ACACCAAGCTTGAAGTGGACCTGCAAGGACTTGCTCCTCAGGCCCTAAGGAGTCCCA 608
Db	541	ACACCAAGCTTGAAGTGGACCTGCAAGGACTTGCTCCTCAGGCCCTAAGGAGTCCCA 600
QY	609	TCTACTGCTGGAGACCTTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGGTCA 668
Db	601	TCTACTGCTGGAGACCTTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCGGTCA 660
QY	669	GCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATACGAAA 728
Db	661	GCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATACGAAA 720
QY	729	AGCCAGCTTGA 740
Db	721	AGCCAGCTTNGA 732

Db 307 GCTGACGGGGCCACGACAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGCGCCCTCC 366
 QY 365 TCAGTCGGCAGCCCAAGCCAGGCTGACATTTATAGACACCGCGCTGCGTTATCGC 424
 Db 367 TCAGTCGGCAGCCCAAGCCAGGCTGACATTTATAGACACCGCGCTGCGTTATCGC 426
 QY 425 GAGGGTCAAAAGTTGAGTGGCTGCTGATGCTGTGTAGCGGAAGTCTTACGGATGA 484
 Db 427 GAGGGTCAAAAGTTGAGTGGCTGCTGATGCTGTGTAGCGGAAGTCTTACGGATGA 486
 QY 485 GCAGTACCAAGGAGTGGGGCGAGCCACACCAACCAAGCAAGATGCGGAAGCTTTCAG 544
 Db 487 GCAGTACCAAGGAGTGGGGCGAGCCACACCAACCAAGCAAGATGCGGAAGCTTTCAG 546
 QY 545 TTTCACACAGCTGGAATCGACCTGCAAGACTTGTCTCTCCAGCGCCCTTAAGGAGTC 604
 Db 547 TTTCACACAGCTGGAATCGACCTGCAAGACTTGTCTCTCCAGCGCCCTTAAGGAGTC 606
 QY 605 CCAGTCTACTGTTGGAGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCG 664
 Db 607 CCAGTCTACTGTTGGAGACCTGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCG 666
 QY 665 GTCAGCCCTGGAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAC 724
 Db 667 GTCAGCCCTGGAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAC 726
 QY 725 GAAAGCCAGCTTGAA 740
 Db 727 GAAAGCCAGCTTGNA 742

RESULT 3
 BM910918
 LOCUS
 DEFINITION AGENCOURT_6615922 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454497
 5', mRNA sequence.
 BM910918
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 BM910918
 767 bp mRNA linear EST 12-MAR-2002
 AGENCOURT_6615922 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454497
 5', mRNA sequence.
 BM910918
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-i@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1949 row: 1 column: 18
 High quality sequence stop: 641.
 Location/Qualifiers

FEATURES
 source
 1. .767
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5454497"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 98"
 /notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 163 a 241 c 237 g 124 t 2 others
 ORIGIN
 Query Match 97.6%; Score 722; DB 12; Length 767;
 Best Local Similarity 99.7%; Fred. No. 3.8e-152;
 Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 7 CGGCTGACGCGGGGTGAGCGCGGCGGCGGATCTTGGAGCCATGGGCGCGCGC 66
 Db 17 CGGCTGACGCGGGGTGAGCGCGGCGGCGGATCTTGGAGCCATGGGCGCGCGC 76
 QY 67 GCGACGCCATCCTGGATCGCTGGAGAACTTACCGCCGAGAGTCAAGAAGTTCAAG 126
 Db 77 GCGACGCCATCCTGGATCGCTGGAGAACTTACCGCCGAGAGTCAAGAAGTTCAAG 136
 QY 127 TGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
 Db 137 TGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196
 QY 187 TGTCCATGAGCGCTTGGACCTTACCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
 Db 197 TGTCCATGAGCGCTTGGACCTTACCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
 QY 247 GCGCGGAGCTACCGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
 Db 257 GCGCGGAGCTACCGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
 QY 307 TGCAGGCGGCGACGACGAGGCTTGGAGCGCGCGGATCTGCGGATCTGCGGATCTGCGGATCT 366
 Db 317 TGCAGGCGGCGACGACGAGGCTTGGAGCGCGCGGATCTGCGGATCTGCGGATCTGCGGATCT 376
 QY 367 AGTCGCGAGCCAGGAGCGGCTGCTGCTTATAGACAGCAGCGGCTGCTGCTTATCGCA 426
 Db 377 AGTCGCGAGCCAGGAGCGGCTGCTGCTTATAGACAGCAGCGGCTGCTGCTTATCGCA 436
 QY 427 GGTTCACAAAAGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
 Db 437 GGTTCACAAAAGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
 QY 487 AGTACAGGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAGTT 546
 Db 497 AGTACAGGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAGTT 556
 QY 547 TCACACGAGCTGGAAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
 Db 557 TCACACGAGCTGGAAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
 QY 607 AGTCTTACTGCTGGAGGAGCTGAGGAGCTGAGGCTGCTTCCAGCAACACTCCCGT 666
 Db 617 AGTCTTACTGCTGGAGGAGCTGAGGAGCTGAGGCTGCTTCCAGCAACACTCCCGT 676
 QY 667 CAGCCCTTGGCAATCCCAACCAATCATCTTCAATCTGATCTTTTATACAAATATACG 725
 Db 677 CAGCCCTTGGCAATCCCAACCAATCATCTTCAATCTGATCTTTTATACAAATATACG 736
 QY 726 AAAAGCCAGCTTGAA 740
 Db 737 AAAAGCCAGCTTGAA 751
 RESULT 4
 BM549665
 LOCUS
 DEFINITION AGENCOURT_6544142 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745741
 5', mRNA sequence.
 BM549665
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE		Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 792)	
TITLE		NIH-MGC http://mgi.nci.nih.gov/	
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT		Unpublished	
CONTACT		Contact: Robert Strausberg, Ph.D.	
EMAIL		Email: cgapbs-remail.nih.gov	
TISSUE		Tissue Procurement: Life Technologies, Inc.	
CDNA		CDNA Library Preparation: Life Technologies, Inc.	
DNA		DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
CLONE		Cloning Distribution: MGC clone distribution information can be	
FOUND		found through the I.M.A.G.E. Consortium/LLNL at:	
HTTP		http://image.llnl.gov	
PLATE		Plate: LLML12769 row: c column: 22	
HIGH		High quality sequence stop: 570.	
FEATURES		Location/Qualifiers	
SOURCE		1..792	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5745741"	
		/tissue_type="leukocyte"	
		/lab_host="DH10B"	
		/clone_lib="NIH MGC 118"	
		/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV	
		(destroyed); RNA source leukocytes from anonymous pool of	
		non-activated adult donors. Library is oligo-dT primed	
		and directionally cloned (EcoRV site is destroyed upon	
		cloning). Average insert size 1.7 kb, insert size range	
		1.2-3.3 kb. Library is normalized and enriched for	
		full-length clones and was constructed by C. Gruber	
		(Invitrogen). Research Genetics tracking code 027. Note:	
		this is a NIH MGC library."	
BASE COUNT		182 a 241 c 242 g 122 t 5 others	
ORIGIN			
Query Match		97.5%; Score 721.2; DB 12; Length 792;	
Best Local Similarity		98.8%; Pred. No. 5.7e-152;	
Matches		726; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
QY	6	CGCGTTCAGCGGGTGCAGCGCGCGGCGGATCCTGGAGCCATGGCGCGCGG	65
DB	19	CCTGGATCAGCGGGTGCAGCGCGGCGGATCCTGGAGCCATGGCGCGCGG	78
QY	66	CGCGAGCCATCCTGGATCGCTGGAGAACCTGACCGCGAGAGCTCAAGATTCAAG	125
DB	79	CGCGAGCCATCCTGGATCGCTGGAGAACCTGACCGCGAGAGCTCAAGATTCAAG	138
QY	126	CTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	185
DB	139	CTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	198
QY	186	CTGTCCATGAGCGCTTGGACCTCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT	245
DB	199	CTGTCCATGAGCGCTTGGACCTCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT	258
QY	246	GGCGCGAGCTCACCCTAACCTGCTGCGGACATGGCGCTGCGAGAGATGCGCGGCGAG	305
DB	259	GGCGCGAGCTCACCCTAACCTGCTGCGGACATGGCGCTGCGAGAGATGCGCGGCGAG	318
QY	306	CTGCAGCGCGCAGCAGCAGCGGCTCTGGAGCGCGCGAGCTGGATCCAGGCCCTCTCT	365
DB	319	CTGCAGCGCGCAGCAGCAGCGGCTCTGGAGCGCGCGAGCTGGATCCAGGCCCTCTCT	378
QY	366	CAGTCGCGCAGCAGCAGCGGCTCTATAGACAGCAGCAGCGGCTGCTGCTATCGCG	425
DB	379	CAGTCGCGCAGCAGCAGCGGCTCTATAGACAGCAGCAGCGGCTGCTGCTATCGCG	438
QY	426	AGGTCACAAAAGCTTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	485
DB	439	AGGTCACAAAAGCTTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	498
QY	486	CAGTACACGAGCGAGTGGCGCGGAGCCACCAACCAAGAGATGCGGAAGCTTTCAGT	545
DB	499	CAGTACACGAGCGAGTGGCGCGGAGCCACCAACCAAGAGATGCGGAAGCTTTCAGT	558
QY	546	TTACACACGAGCGTGGAACTGGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGAGTCC	605
DB	559	TTACACACGAGCGTGGAACTGGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGAGTCC	618
QY	606	CAGTCTACCTGCTGGTGGAGGAGCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGG	665
DB	619	CAGTCTACCTGCTGGTGGAGGAGCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGG	678
QY	666	TCAGCCCTGCGCAATCCCAACCAATCATCTGATCTCTTTTATACAAATATACG	725
DB	679	TCAGCCCTGCGCAATCCCAACCAATCATCTGATCTCTTTTATACAAATATACG	738
QY	726	AAAGCCGAGCTTGAA 740	
DB	739	AAAGCCGAGCTTGAA 753	
RESULT 5		750 bp mRNA linear EST 15-MAY-2001	
BG764161		602737190F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862182 5',	
LOCUS		mRNA sequence.	
DEFINITION		BG764161	
ACCESSION		BG764161.1 GI:14074814	
VERSION		EST.	
KEYWORDS		Homo sapiens (human)	
SOURCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
ORGANISM		Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 750)	
AUTHORS		NIH-MGC http://mgi.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished	
COMMENT		Contact: Robert Strausberg, Ph.D.	
		Email: cgapbs-remail.nih.gov	
		Tissue Procurement: ATCC/DCD/DTP	
		cDNA Library Preparation: Ling Hong/Rubin Laboratory	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LLNL at:	
		http://image.llnl.gov	
		Plate: LLML1722 row: d column: 23	
		High quality sequence stop: 721.	
FEATURES		Location/Qualifiers	
SOURCE		1..750	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4862182"	
		/tissue_type="melanotic melanoma, high MDR (cell line)"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH MGC_49"	
		/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:	
		EcoRI; cDNA made by oligo-dT priming. Directionally cloned	
		into EcoRI/XhoI sites using the following 5' adaptor:	
		GGCAGGAG(G). Size-selected >500bp for average insert size	
		1.8kb. Library constructed by Ling Hong in the laboratory	
		of Gerald M. Rubin (University of California, Berkeley)	
		using ZAP-cDNA synthesis kit (Stratagene) and Superscript	
		II RT (Life Technologies). Note: this is a NIH_MGC	
		Library."	
BASE COUNT		165 a 232 c 235 g 118 t	
ORIGIN			
Query Match		96.4%; Score 713.4; DB 12; Length 750;	
Best Local Similarity		99.7%; Pred. No. 3.2e-150;	
Matches		725; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	

15 GCGGGGTGAGCGCGGACGCGCGGGGATCTGAGAGCCATGGGCGCGCGCGACGCC 74
 Db GCGGGGTGAGCGCGGACGCGCGGGGATCTGAGAGCCATGGGCGCGCGCGACGCC 61
 75 ATCTTGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAGCTGAAGCTG 134
 Db ATCTTGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAGCTGAAGCTG 121
 135 CTGTGCGTCCGCTGCGGAGGAGCTACGGGCGCATCCCGCGGGGCGCTGTGTCCATG 194
 Db CTGTGCGTCCGCTGCGGAGGAGCTACGGGCGCATCCCGCGGGGCGCTGTGTCCATG 181
 195 GACGCTTGAGACCTCACCGACAAAGCTGTGCTTACCTGGAGACCTACCGGCGCGGAG 254
 Db GACGCTTGAGACCTCACCGACAAAGCTGTGCTTACCTGGAGACCTACCGGCGCGGAG 241
 255 CTCACGCTAACTGTGCGGACATGGGCTGCGAGGATGCGCGGCGAGCTGCGAGGG 314
 Db CTCACGCTAACTGTGCGGACATGGGCTGCGAGGATGCGCGGCGAGCTGCGAGGG 301
 315 GCCAGCCACAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCAGTCGGCA 374
 Db GCCAGCCACAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTCAGTCGGCA 361
 375 GCGAGCCAGGCTGACCTTTATAGACAGCACCGGCTGCGCTTATCGGAGGCTACA 434
 Db GCGAGCCAGGCTGACCTTTATAGACAGCACCGGCTGCGCTTATCGGAGGCTACA 421
 435 AACGTTGAGTGTGCTGATGCTGTGACGGGAGGCTCTACGGATGAGCAGTACCG 494
 Db AACGTTGAGTGTGCTGATGCTGTGACGGGAGGCTCTACGGATGAGCAGTACCG 481
 495 GCAGTGGCGGCGGAGCCACCAACCAAGAGTGGGAGCTCTTCAAGTTTACACCA 554
 Db GCAGTGGCGGCGGAGCCACCAACCAAGAGTGGGAGCTCTTCAAGTTTACACCA 541
 555 GCGTGAATGACCTGCAAGGACTTGTCTCCAGGCGCTAAGGAGTCCAGTCCCTAC 614
 Db GCGTGAATGACCTGCAAGGACTTGTCTCCAGGCGCTAAGGAGTCCAGTCCCTAC 601
 615 CTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTTCCGAGCACACTCCGGTCAAGCCCT 674
 Db CTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTTCCGAGCACACTCCGGTCAAGCCCT 661
 675 GCGAATCCCAACCAATCATCTGAATCTGATCTTTTATATACA-CAATATACGAAGAAGCA 733
 Db GCGAATCCCAACCAATCATCTGAATCTGATCTTTTATATACGAATATACGAAGAAGCA 721
 734 GCTTGA 740
 Db GCTTGA 728

RESULT 6
 CA309883/c
 LOCUS
 DEFINITION UI-H-Ftl1-bid-e-07-0-UI.s1 NCI CGAP Ftl1 Homo sapiens cDNA clone
 UI-H-Ftl1-bid-e-07-0-UI 3', mRNA sequence.
 CA309883
 VERSION 1 (bases 1 to 725)
 KEYWORDS EST.
 SOURCE CA309883.1 GI:24472937
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 725)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-Ftl1-bid-e-07-0-UI"
 /tissue_type="Aveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Ftl1"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site1: EcoR I; Site2: Not I; NCI CGAP Ftl1 is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (drl)18 tail. The sequence tag for this library is GGCATGCCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
 TAG LIB=UI-H-Ftl1
 TAG_TISSUE=Human Lung Aveolar Macrophage
 TAG_SEQ=GGCCATGCCG"

BASE COUNT 119 a 221 c 226 g 159 t

Query Match 95.2%; Score 704.8; DB 14; Length 725;
 Best Local Similarity 99.7%; Pred. No. 2.7e-148;
 Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 GCGGCGGGGATCTCGAGCCATGGGCGCGCGCGCGCATCTCTGATCGCTGGAG 92
 Db 724 GCGGCGGGGATCTCTGAGCCATGGGCGCGCGCGCGCATCTCTGATCGCTGGAG 665
 QY 93 AACCTGACCGCGGAGAGCTCAAGAGTTCAAGCTGAAGCTGCTGTGGTGGCTGGCC 152
 Db 664 AACCTGACCGCGGAGAGCTCAAGAGTTCAAGCTGAAGCTGCTGTGGTGGCTGGCC 605
 QY 153 GAGGCTACGGGCGCATCCCGGGGCGCTGCTGCTCATGAGCCCTTGGACCTCAC 212
 Db 604 GAGGCTACGGGCGCATCCCGGGGCGCTGCTGCTCATGAGCCCTTGGACCTCAC 545
 QY 213 GACAACTGCTCAGCTTCTTACTGTGAGACCTACCGCGCGCGAGCTCACCGTAACTGCTG 272
 Db 544 GACAACTGCTCAGCTTCTTACTGTGAGACCTACCGCGCGCGAGCTCACCGTAACTGCTG 485
 QY 273 GCGCATGCGGCTCGAGAGATGCGCGGAGCTGTCAGGCGGCGCACCGCACCGAGGCTCT 332
 Db 484 GCGCATGCGGCTCGAGAGATGCGCGGAGCTGTCAGGCGGCGCATGTCAGGCGGCTCT 425
 QY 333 GAGGCGGCGGAGCTGGGATCCAGGCGGCTCTCTGCTGCGGAGCCAGCCAGGCGCTGCAC 392
 Db 424 GAGGCGGCGGAGCTGGGATCCAGGCGGCTCTCTGCTGCGGAGCCAGCCAGGCGCTGCAC 365
 QY 393 TTTATAGACCAACACCGGCTGCGCTTATCGGAGGCTCACAAAGCTTGAGTGGCTGCTG 452
 Db 364 TTTATAGACCAACACCGGCTGCGCTTATCGGAGGCTCACAAAGCTTGAGTGGCTGCTG 305
 QY 453 GATGCTCTGTACGGAAAGTCTCTGACGATGAGCAGTACCGAGGAGTGTGGGCCGAGGCC 512

```

Db      304  GATGCTCTGTACCGGAGGCTCTTCACGAGTACAGAGTACAGGAGCTGGGGCCGAGCC 245
QY      513  ACCAACCCAGCAAGATGCGGAGCTCTTCAGTTTCACACCCAGCTGGAACCTGACCTGC 572
Db      244  ACCAACCCAGCAAGATGCGGAGCTCTTCAGTTTCACACCCAGCTGGAACCTGACCTGC 185
QY      573  AAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCAGTCTCTACCTGTGTGAGGACCTGGAG 632.
Db      184  AAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCAGTCTCTACCTGTGTGAGGACCTGGAG 125
QY      633  CGAGCTGAGGCTCTCTCCAGCAACACCTCCGCTCAGCCCTGGCAATCCACCAAATCA 692
Db      124  CGAGCTGAGGCTCTCTCCAGCAACACCTCCGCTCAGCCCTGGCAATCCACCAAATCA 65
QY      693  TCCTGAATCTGATCTTTTATACACATATACGAAAGCAGCTTGAA 740
Db      64  TCCTGAATCTGATCTTTTATACACATATACGAAAGCAGCTTGAA 17

RESULT 7
LOCUS   BG767422
DEFINITION 602741219F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4871020 5',
mRNA sequence.
ACCESSION BG767422
VERSION   BG767422.1 GI:114078075
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC/DCTD/BNP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1745 row: e column: 05
High quality sequence stop: 696.

FEATURES
source
1..856
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4871020"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 49"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using 2AP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 194 a 234 c 279 g 149 t
ORIGIN
Query Match 93.4%; Score 690.8; DB 12; Length 856;
Best Local Similarity 99.6%; Pred. No. 4e-145; 2; Gaps 1;
Matches 703; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 15 GCGGGGTGAGCGCGGCGGCGGAGTCTTGAGCCATGGGCGCGCGCGACGCC 74
|||||

```

```

Db      2  GCGGGGTGAGCGCGGCGGCGGATCCTGAGCCATGGGGCGCGCGACGCC 61
QY      75  ATCCCTGGATGGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 134
Db      62  ATCCCTGGATGGCTGGAGAACCTGACCCGCGAGGAGCTCAAGAGTTCAAGCTGAAGCTG 121
QY      135  CTGTCCGTGCCGCTCGCGGAGGGCTACCGGGGCGATCCCGCGGGGCGCGCTGTGTCCATG 194
Db      122  CTGTCCGTGCCGCTCGCGGAGGGCTACCGGGGCGATCCCGCGGGGCGCGCTGTGTCCATG 181
QY      195  GACGCTTTGGACCTCACCGACAAGCTGTGTACCTTCTACCTGGAGACCTACCGGCGCCAG 254
Db      182  GACGCTTTGGACCTCACCGACAAGCTGTGTACCTTCTACCTGGAGACCTACCGGCGCCAG 241
QY      255  CTCACCGCTAAGCTGCTCGCGACATGGCCCTGCAAGGAGATGGCGGCGAGCTGCAGCGC 314
Db      242  CTCACCGCTAAGCTGCTCGCGACATGGCCCTGCAAGGAGATGGCGGCGAGCTGCAGCGC 301
QY      315  GCCACGACCAAGGGCTCTGGAGCGCGCGAGTGGGATCCAGGCCCTCTCAGTCCGCA 374
Db      302  GCCACGACCAAGGGCTCTGGAGCGCGCGAGTGGGATCCAGGCCCTCTCAGTCCGCA 361
QY      375  GCCAAGCAGGCTGCATTTATAGACCAGCACCGGGCTGGCTTATCGGAGGTCACA 434
Db      362  GCCAAGCAGGCTGCATTTATAGACCAGCACCGGGCTGGCTTATCGGAGGTCACA 421
QY      435  RACGTTGAGTGGCTGCTGGATGCTGTACGGGAGGTCCTGACGAGTACGAGTACACAG 494
Db      422  RACGTTGAGTGGCTGCTGGATGCTGTACGGGAGGTCCTGACGAGTACGAGTACACAG 481
QY      495  GCAGTGGCGGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTACACCA 554
Db      482  GCAGTGGCGGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTACACCA 541
QY      555  GCCTGGAATCGAAGCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGAGTCCAGTCTAC 614
Db      542  GCCTGGAATCGAAGCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGAGTCCAGTCTAC 601
QY      615  CTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCCAGCAACACTCCGCTCAGCCCT 674
Db      602  CTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGCTCAGCCCT 661
QY      675  GGC-AATCCCAACCAATCATCTGAATCTGATCTTTTATACACAA 719
Db      662  GGCATATCCCAACCAATCATCTGAATCTGATCTTTTATACACAA 707

RESULT 8
LOCUS   BG337806
DEFINITION 602435709F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4553378 5',
mRNA sequence.
ACCESSION BG337806
VERSION   BG337806.1 GI:13144342
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 748)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1251 row: b column: 03

```


High quality sequence stop: 710.
 Location/Qualifiers
 1. 748
 /organism="Homo sapiens"
 /mol_type="mrna"
 /db_xref="taxon:9606"
 /clone="IMAGE:4553378"
 /tissue_type="leiomyosarcoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_46"
 /note="Organ: uterus; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 153 a 239 c 238 g 118 t
 ORIGIN
 Query Match 93.1%; Score 688.6; DB 10; Length 748;
 Best Local Similarity 98.5%; Pred No. 1.2e-144;
 Matches 716; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
 QY 15 GCGGGGTGAGCGGCGAGCGCGGGGATCTCTGAGCCATGGGGCGCGCGCGAGCC 74
 Db 2 GCGGGGTGAGCGGCGAGCGCGGGGATCTCTGAGCCATGGGGCGCGCGCGAGCC 61
 QY 75 ATCTTGATGGCTGGAGAACCTTACCGCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTG 134
 Db 62 ATCTTGATGGCTGGAGAACCTTACCGCGAGGAGCTCAAGAAGTTCAAGCTGAAGCTG 121
 QY 135 CTGTCGTCGCTGCGGAGGGCTACGGCGCATCCGCGGGCGCGCTGCTGTCCTCATG 194
 Db 122 CTGTCGTCGCTGCGGAGGGCTACGGGCGATCCCGGGGCGCGCTGCTGTCCTCATG 181
 QY 195 GAGCCTTGGACCTCACGCAAGAGCTGCTCAGCTTTCTACCTGGAGACTACGGGCGCGAG 254
 Db 182 GAGCCTTGGACCTCACGCAAGCTGCTGCTGCTTCTACCTGGAGACTACGGGCGCGAG 241
 QY 255 CTCACCGCTAACGTCGTCGCGCATGGGCTTGAGGAGATGGCGGGCGAGCTGACGGCG 314
 Db 242 CTCACCGCTAACGTCGTCGCGCATGGGCTTGAGGAGATGGCGGGCGAGCTGACGGCG 301
 QY 315 GCCACGACACAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCAGTCGGCA 374
 Db 302 GACACGACACAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCAGTCGGCA 361
 QY 375 GCCAAGCCAGGCTCTGACTTTATAGACAGACCGCGGCTGGCTTATCGGAGGGTCACA 434
 Db 362 GCCAAGCCAGGCTCTGACTTTATAGACAGACCGCGGCTGGCTTATCGGAGGGTCACA 421
 QY 435 AACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGACAGTACCA 494
 Db 422 AACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGACAGTACCA 481
 QY 495 GCAGTCGGGCGGAGCCACCAACCCAGACAGATGCGGAAGCTCTTCAGTTTCACACCA 554
 Db 482 GCAGTCGGGCGGAGCCACCAACCCAGACAGATGCGGAAGCTCTTCAGTTTCACACCA 541
 QY 555 GCCTGGAATGGACCTGCAAGGACTTGCTCTCTCCAGGCCCTTAAGGGAGTCCCACTCTAC 614
 Db 542 GCCTGGAATGGACCTGCAAGGACTTGCTCTCTCCAGGGCTTAGGGAGTCCCACTCTAC 601
 QY 615 CTGCTGGAGGACTTGGAGCGGAGTGGGCTCTTCTCCAGCAACACTCCGGTCAGCCCC- 673
 Db 602 CTGCTGGAGGACTTGGAGCGGAGTGGGCTCTTCTCCAGCAACACTCCGGTCAGCCCC 661
 QY 674 TGGCAATCCACAAATCATCTGAACTGATCTTTTATACAAATATAGCAAAAGCCA 733
 Db 662 TGTCAATCCAGCAGATCATCTGAACTGTATC-TTATATACAAATATAGCAAGCCA 720

309 CAGGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGATCCAGCCCTCTCTCAG 368
 438 CAGGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGATCCAGCCCTCTCTCAG 379
 369 TCGGACGCAAGCAGGCTGCACTTTATAGACGACGCGGCTCGCTTATCGGAGG 428
 378 TCGGACGCAAGCAGGCTGCACTTTATAGACGACGCGGCTCGCTTATCGGAGG 319
 429 GTACACAACTTGAAGTGGTCTGGATGCTCTGTAGGGAAGTCTGTAGGATGAGCAG 488
 318 GTACACAACTTGAAGTGGTCTGGATGCTCTGTAGGGAAGTCTGTAGGATGAGCAG 259
 489 TACGAGCAGTGGGCGCGCCAGCCACCAACCAAGCAAGATCGGAAGCTCTTCAGTTTC 548
 258 TACGAGCAGTGGGCGCGCCAGCCACCAACCAAGCAAGATCGGAAGCTCTTCAGTTTC 199
 549 ACACGAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCAGGCCCTAAAGGAGTCCAG 608
 198 ACACGAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCAGGCCCTAAAGGAGTCCAG 139
 609 TCTTACTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGTCA 668
 138 TCTTACTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGTCA 79
 669 GCGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGAAA 728
 78 GCGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGAAA 19
 729 AGCCAGCTTGAA 740
 18 AGCCAGCTTGAA 7

RESULT 10
 BM051141
 LOCUS
 DEFINITION
 603634063F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:5424590 5', mRNA sequence.
 741 bp mRNA linear EST 07-NOV-2001
 BM051141
 GI:16780408
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 741)
 NIH-MGC http://img.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ARCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM1885 row: n column: 15
 High quality sequence stop: 643.
 Location/Qualifiers
 1. .741
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5424590"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_43"
 /notes="Organ: eye; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong

FEATURES
 Source

BASE COUNT 139 a 243 c 237 g 122 t
 ORIGIN
 Query Match 90.2%; Score 667.6; DB 12; Length 741;
 Best Local Similarity 97.1%; Pred. No. 6.4e-140;
 Matches 701; Conservative 0; Mismatches 19; Indels 2; Gaps 2;
 QY 7 CGGCTTCAGCGGGGTGAGCGGCGGCGGAGTCTTGGAGCCATGGGGCGCGCGC 66
 Db 21 CGGCTTCAGCGGGGTGAGCGGCGGCGGAGTCTTGGAGCCATGGGGCGCGCGC 80
 QY 67 GGGAGCCCATCTCGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGC 126
 Db 81 GCGAGCCCATCTCGATGGCTGGAGAACCTGACCGCGGAGTCTCAAGAGTTCAAGC 140
 QY 127 TGAAGCTGCTGTCGGTGGCTGCGCGAGGGCTACGGGCGCATCCCGGGGCGGCTGC 186
 Db 141 TGAAGCTGCTGTCGGTGGCTGCGCGAGGGCTACGGGCGCATCCCGGGGCGGCTGC 200
 QY 187 TGTCCATGGAGCCCTTGGACCTTACCGAAGAGTGGTCACTTCTTACCTGGAGAGCTACG 246
 Db 201 TGTCCATGGAGCCCTTGGACCTTACCGAAGAGTGGTCACTTCTTACCTGGAGAGCTACG 260
 QY 247 GGGCGGAGCTACCGCTTAACGCTGCGGAGATGGGCTCGAGAGATGGCGGGGCGC 306
 Db 261 GCGCGGAGCTACCGCTTAACGCTGCGGAGATGGGCTCGAGAGATGGCGGGGCGC 320
 QY 307 TGCAGGCGGCGACGACACAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTC 366
 Db 321 TGCAGGCGGCGACGACACAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTC 380
 QY 367 AGTCGGGAGCCCAAGCAGGCTTGCACTTTATAGACCAAGCCGCGCTGCGTTATCGGA 426
 Db 381 AGTCGGGAGCCCAAGCAGGCTTGCACTTTATAGACCAAGCCGCGCTGCGTTATCGGA 440
 QY 427 GGGTCACAAAGCTTGAAGTGGCTGCTGTAGGGAAGTCTCGAGGATGAGC 486
 Db 441 GGGTCACAAAGCTTGAAGTGGCTGCTGTAGGGAAGTCTCGAGGATGAGC 500
 QY 487 AGTACAGGAGTGGCGGCGGAGCCACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
 Db 501 AGTACAGGAGTGGCGGCGGAGCCCATATACCAAGCAAGATGCGGAAGCTCTTCAGTT 560
 QY 547 TCACACGAGCCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAGGGAGTCCC 606
 Db 561 TCACACGAGCCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAGGGAGTCCC 620
 QY 607 AGTCCTACTGGTGGAGGACCTTGGAGGAGTGGAGTCTTCCCGAGCAACATCCGG 665
 Db 621 AGTCCTACTGGTGGAGGACCTTGGAGGAGTGGAGTCTTCCCGAGCAACATCCGG 680
 QY 666 TCAGGCGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGC 725
 Db 681 TCAGGCGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAGC 739
 726 AA 727
 740 AA 741
 RESULT 11
 BM972785/c
 LOCUS
 DEFINITION
 UI-CF-EC1-abr-k-24-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
 UI-CF-EC1-abr-k-24-0-UI 3', mRNA sequence.
 BM972785
 GI:19590376
 VERSION
 KEYWORDS
 SOURCE
 Homo sapiens (human)

in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 676)
JOURNAL Normalization and subtraction: two approaches to facilitate gene
MEDLINE discovery
PUBMED Genome Res. 6 (9), 791-806 (1996)
COMMENT 97044477
8889548
Contact: McCray, PB
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             Location/Qualifiers
     source            1..676
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="UI-CF-ECl-abr-k-24-0-UI"
                        /tissue_type="Lung"
                        /dev_stage="Adult and Fetal"
                        /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                        /clone_lib="UI-CF-ECl"
                        /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
                        modified polylinker; Site 1: EcoR I; Site 2: Not I;
                        UI-CF-ECl is a normalized cDNA library containing the
                        following tissue(s): Normal lung from adult and from fetal
                        day 64, day 87, week 19 and week 42. The library was
                        constructed according to Bonaldo, Lennon and Soares,
                        Genome Research, 6:791-806, 1996. First strand cDNA
                        synthesis was primed with an oligo-dT primer containing a
                        Not I site. Double stranded cDNA was ligated to an EcoR I
                        adaptor, digested with Not I, and cloned directionally
                        into pT73-Pac vector. The oligonucleotide used to prime
                        the synthesis of first-strand cDNA contains a library tag
                        sequence that is located between the Not I site and the
                        (dT)18 tail. The sequence tag for this library is
                        AAGTGCTTAC.
                        TAG_LIB=UI-CF-ECl
                        TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
                        and 380-383
                        TAG_SEQ=AAGTGCTTAC"

BASE COUNT 112 a 199 c 211 g 154 t
ORIGIN

Query Match      89.2%; Score 660; DB 12; Length 676;
Best Local Similarity 100.0%; Pred. No. 3.2e-138;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GATGCCGTGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGTGTCG 140
Db |||||||
QY 676 GATGCCGTGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGCTGAAGCTGTGTCG 617
Db |||||||
QY 141 GTGCCCGTGGCGAGGAGCTACGGCGGCATCCCGCGGGCGCGCTGCTGCCATGGACGCC 200
Db |||||||
QY 616 GTGCCCGTGGCGAGGAGCTACGGCGGCATCCCGCGGGCGCGCTGCTGCCATGGACGCC 557
Db |||||||
QY 201 TTGAGACTCACCGACAAGTGTGTAGTCTTACTCGAGACCTACCGCGCGGAGTCTACC 460
Db |||||||
QY 556 TTGAGACTCACCGACAAGTGTGTAGTCTTACTCGAGACCTACCGCGCGGAGTCTACC 497
Db |||||||

```

```

QY 261 GCTAACGTGCTGCGGCACATGGGCTTGAGAGATGGCCGGGAGCTTCAGCGGCGCCACG 320
Db |||||||
QY 496 GCTAACGTGCTGCGGCACATGGGCTTGAGAGATGGCCGGGAGCTTCAGCGGCGCCACG 437
Db |||||||
QY 321 CACCAGGGCTCTGAGCGCGCCAGCTGGGATCCAGGCGCCCTCTCTCAGTCGCGAGCCCAAG 380
Db |||||||
QY 436 CACCAGGGCTCTGAGCGCGCCAGCTGGGATCCAGGCGCCCTCTCTCAGTCGCGAGCCCAAG 377
Db |||||||
QY 381 CCAGGCTCTGCACTTTATAGACACAGCACCGGCTGCGCTTATCGCGAGGCTCACAAACGTT 440
Db |||||||
QY 376 CCAGGCTCTGCACTTTATAGACACAGCACCGGCTGCGCTTATCGCGAGGCTCACAAACGTT 317
Db |||||||
QY 441 GAGTGGCTGCTGGATGCTCTGTATCGGGAAGTCTCTGACGGATGACAGTACAGGACGTG 500
Db |||||||
QY 316 GAGTGGCTGCTGGATGCTCTGTATCGGGAAGTCTCTGACGGATGACAGTACAGGACGTG 257
Db |||||||
QY 501 CGGCGCGAGCCCAACCAAGCAAGATGGGAGCTCTTCAGTTTCACACCGACCTGG 560
Db |||||||
QY 256 CGGCGCGAGCCCAACCAAGCAAGATGGGAGCTCTTCAGTTTCACACCGACCTGG 197
Db |||||||
QY 561 AACTGAGCTCTCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCAGTCTCTACCTGGTG 620
Db |||||||
QY 196 AACTGAGCTCTCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCAGTCTCTACCTGGTG 137
Db |||||||
QY 621 GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTCAGCCCTGGCAAT 680
Db |||||||
QY 136 GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTCAGCCCTGGCAAT 77
Db |||||||
QY 681 CCCACCAATCATCTGATCTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 740
Db |||||||
QY 76 CCCACCAATCATCTGATCTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 17
Db |||||||

```

```

RESULT 12
BE908204 LOCUS
DEFINITION BE908204 675 bp mRNA linear EST 20-OCT-2000
mRNA sequence.
ACCESION BE908204.1 GI:10402543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 675)
AUTHORS NIH-MGC http://mgc.hci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9704 row: j column: 16
High quality sequence stop: 660.
Location/Qualifiers
1..675
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3902319"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life

```

Technologies."

BASE COUNT	151 a	214 c	198 g	112 t
ORIGIN				
Query Match	88.4%; Score 654.2; DB 10; Length 675;			
Best Local Similarity	99.5%; Pred. No. 6.5e-137;			
Matches	656; Conservative	0; Mismatches	3; Indels	0; Gaps
QY	82	ATGCGCTGAGAACCTGACCGCGGAGGCTCAAGAGTTCAAGCTGAAGCTGCTGTCGG	141	
Db	1	ATGCGCTGAGAACCTGACCGCGGAGGCTCAAGAGTTCAAGCTGAAGCTGCTGTCGG	60	
QY	142	TGCGCTGCGCGAGGCTACGCGCGATCCCGCGGCGCGCTGCTGCTCCATGAGCGCT	201	
Db	61	TGCGCTGCGCGAGGCTACGCGCGATCCCGCGGCGCGCTGCTGCTCCATGAGCGCT	120	
QY	202	TGACCTACCGAACAGCTGCTGAGCTTCTACTTGAGACCTACGCGCGAGCTCACCG	261	
Db	121	TGACCTACCGAACAGCTGCTGAGCTTCTACTTGAGACCTACGCGCGAGCTCACCG	180	
QY	262	CTAACGTGCTGCGGACATGGGCTGCGAGAGTGGCGGCGAGCTGCGAGGCGGCGACG	321	
Db	181	CTAACGTGCTGCGGACATGGGCTGCGAGAGTGGCGGCGAGCTGCGAGGCGGCGACG	240	
QY	322	ACCAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCTCTCAGTCCGCGAGCAAGC	381	
Db	241	ACCAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCGCCCTCTCTCAGTCCGCGAGCAAGC	300	
QY	382	CAGGCTGCACTTTATAGACAGACCGGGCTGCGCTTATCGGAGGTCACAAAGTTG	441	
Db	301	CAGGCTGCACTTTATAGACAGACCGGGCTGCGCTTATCGGAGGTCACAAAGTTG	360	
QY	442	AGTGGCTGCTGAGTCTGTGAGGAGTCTGACGAGTCCGCGAGTACGAGGAGTGC	501	
Db	361	AGTGGCTGCTGAGTCTGTGAGGAGTCTGACGAGTCCGCGAGTACGAGGAGTGC	420	
QY	502	GGGCGGAGCCCAACCAAGCAAGATCGGAGGCTCTTCAAGTTTCAACAGGCTTGA	561	
Db	421	GGGCGGAGCCCAACCAAGCAAGATCGGAGGCTCTTCAAGTTTCAACAGGCTTGA	480	
QY	562	ACTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGAGTCCAGTCTTACCTGTTGG	621	
Db	481	ACTGACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGAGTCCAGTCTTACCTGTTGG	540	
QY	622	AGGACCTGAGCGGAGCTGAGGCTCTCCAGCACACTCCGCTCAGCCCTGCGCAATC	681	
Db	541	AGGACCTGAGCGGAGCTGAGGCTCTCCAGCACACTCCGCTCAGCCCTGCGCAATC	600	
QY	682	CCACCAATCATCTGAACTGATCTTTTATACAAATATACGAAAGCCAGCTTGA	740	
Db	601	CCACCAATCATCTGAACTGATCTTTTATACAAATATACGAAAGCCAGCTTGA	659	

RESULT 13
 LOCUS BG255521
 DEFINITION 602367671F1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:4476088 5', mRNA linear EST 13-PEB-2001
 mRNA sequence.
 ACCESSION BG255521
 VERSION BG255521.1 GI:12765259
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 744)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTP/DRP
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10302 row: m column: 17
 High quality sequence stop: 661.
 Location/Qualifiers

1..744
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4476088"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 159 a 227 c 238 g 120 t

Query Match	87.5%; Score 647.4; DB 10; Length 744;
Best Local Similarity	97.0%; Pred. No. 2.2e-135;
Matches	713; Conservative 0; Mismatches 16; Indels 6; Gaps 5;
QY	8 GGCTGACGGGGTGAAGCGGCGGAGGATCCTGGAGCCATGGGCGGCGCGG 67
Db	1 GGCTGACGGGGTGAAGCGGCGGAGGATCCTGGAGCCATGGGCGGCGCGG 60
QY	68 CGAGCCATCTTGGATGGCTGGAGAACTCAGCCGCGAGGAGCTCAAGAGTTCAAGCT 127
Db	61 CGAGCCATCTTGGATGGCTGGAGAACTCAGCCGCGAGGAGCTCAAGAGTTCAAGCT 120
QY	128 GAAGTGTGTGTCGGTGGCGGCTGCGGAGGCTACGGGCGCATCCGCGGGGCGGCTGCT 187
Db	121 GAAGTGTGTGTCGGTGGCGGCTGCGGAGGCTACGGGCGCATCCGCGGGGCGGCTGCT 180
QY	188 GTCCATGAGCGCTTGGACCTCAGGACCAAGCTGGTACGCTTCTACCTGGAGACCTACGG 247
Db	181 GTCCATGAGCGCTTGGACCTCAGGACCAAGCTGGTACGCTTCTACCTGGAGACCTACGG 240
QY	248 CGCGAGCTCACCGCTACGCTGCGGACATGGGCTGCGAGAGATGGCGGCGAGCT 307
Db	241 CGCGAGCTCACCGCTACGCTGCGGACATGGGCTGCGAGAGATGGCGGCGAGCT 300
QY	308 GCAGCGGCCACGACACGAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCA 367
Db	301 GCAGCGGCCACGACACGAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCA 360
QY	368 GTGGCAGCCAAAGCAGCGCTTGCACTTTATAGACAGCAGCGGCTGGGCTTATCGGAG 427
Db	361 GTGGCAGCCAAAGCAGCGCTTGCACTTTATAGACAGCAGCGGCTGGGCTTATCGGAG 420
QY	428 GGTCAACAACTTGAAGTGGCTGCTGCTGTGAGGAAAGTCTTACGAGATGAGCA 487
Db	421 GGTCAACAACTTGAAGTGGCTGCTGCTGTGAGGAAAGTCTTACGAGATGAGCA 480
QY	488 GTACCCAGGCTGGGGCGGAGCCACCAACCCAGCAAGATGGGAAAGCTCTTCAGTTT 547
Db	481 GTACCCAGGCTGGGGCGGAGCCACCAACCCAGCAAGATGGGAAAGCTCTTCAGTTT 539
QY	548 CACACGAGCTGGAACCTGGAAGCTTGTCTCTCCAGGCCCTTAAGGAGTCCCA 607
Db	540 CACACGAGCTT-GAAGTGAAGCTTGTCTCTCCAGGCCCTTAAGGAGTCCCA 597
QY	608 GTCTTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTC 667
Db	598 GTCTTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCT-CTTCAGGCAACACTCCGCTC 656
QY	668 AGCCCTGGCAATCCCAACCAATCATCTCTGAAT--CTGATCTTTTATACAAATATACG 725

```

Db      657  AGCCCCGGGAATCCACAAAATATCTTGGATTCTGATCTTTTATACAAATATACG 716
QY      726  AAAAGCCAGCTTGAA 740
Db      717  AAAAGCCAGTTGAAA 731

RESULT 14
BE909218
LOCUS   BE909218
DEFINITION BE909218
ACCESSION BE909218
VERSION   BE909218.1 GI:10404770
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 687)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9707 row: k column: 19
High quality sequence stop: 673.
Location/Qualifiers
1. .638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3903498"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 140 a 219 c 214 g 114 t
ORIGIN

Query Match 86.4%; Score 639; DB 10; Length 687;
Best Local Similarity 98.7%; Pred. No. 1.7e-133;
Matches 676; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 48 GGAGCCATGGGGCGCG-CGGCGACGCCATCTCTGGATGGCTGGAGAACTGACCCCGCA 106
Db 1 GGAAGCCATGGGGCGCGCGCGCGACGCCATCTCTGGATGGCTGGAGAACTGACCCCGCA 60

QY 107 GGAGCTCAAGAGTTCAAGCTGAAGCTGTGTGCTGCGTCCGCTGCGCGAGGGCTACGGGCG 166
Db 61 GGAGCTCAAGAGTTCAAGCTGAAGCTGTGTGCTGCGTCCGCTGCGCGAGGGCTACGGGCG 120

QY 167 CATCCCGGGGGCGCGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTGTAG 226
Db 121 CATCCCGGGGGCGCGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTGTAG 180

QY 227 TTCTTACCTGGAGACTACGGCGCCGAGCTCACCGCTTAAGCTGCTCGCGACATGGGCT 286
Db 181 TTCTTACCTGGAGACTACGGCGCCGAGCTCACCGCTTAAGCTGCTCGCGACATGGGCT 240

QY 287 GCAGGAGATGGCCGGCGAGCTGCAGGCGGCGACGCCACCGAGGCTCTGGAGCGCGCCAGC 346

```

```

Db      241  GCAGGAGATGGCGGGCAGCTGCAGGCGGCCACGCACAGGGCTCTGGAGCGGCCAGC 300
QY      347  TGGGATCCAGGCCCTCTCTCAGTCCGCGACCAAGCCAGCCCTGCACTTTATAGACCAGCA 406
Db      301  TGGGATCCAGGCCCTCTCTCAGTCCGCGACCAAGCCAGCCCTGCACTTTATAGACCAGCA 360
QY      407  CCGGCTCGGCTTATCGCGAGGGTCAAAAGTTGAGTGGCTGCTGGATGCTCTGTACGG 466
Db      361  CCGGCTCGGCTTATCGCGAGGGTCAAAAGTTGAGTGGCTGCTGGATGCTCTGTACGG 420
QY      467  GAAGTCTCTGCGGATGAGCAGTACAGGCGAGTCCGCGCCGAGCCCAACCAAGCA 526
Db      421  GAAGTCTCTGCGGATGAGCAGTACAGGCGAGTCCGCGCCGAGCCCAACCAAGCA 480
QY      527  GATGCGAAGCTCTTTCAGTTTTCACACCAAGCCCTGGAACCTGGACCTCAAGGAGTTGCTCT 586
Db      481  GATGCGAAGCTCTTTCAGTTTTCACACCAAGCCCTGGAACCTGGACCTCAAGGAGTTGCTCT 540
QY      587  CCAGGCCCTAAGGGAGTCCCAAGTCTTCTGCTGGTGGAGACCTCGAGCGGAGCTGAGGCTC 646
Db      541  CCAGGCCCTAAGGGAGTCCCAAGTCTTCTGCTGGTGGAGACCTCGAGCGGAGCTGAGGCTC 600
QY      647  CTTCCAGCAACACTCCGCTCAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATC 706
Db      601  --TTCCAGCACACTCCGCTCAGCCCTGG--AATCCCAACCAATCATCTGAATCTGATC 657
QY      707  TTTTATACAAATATACGAAAAGC 731
Db      658  TTTTATACAAATATACGAAAAGC 682

RESULT 15
BM998354/c
LOCUS   BM998354
DEFINITION UI-H-DT1-3wc-p-21-0-UI.s1 NCI_CGAP_DTI Homo sapiens cDNA clone
IMAGE:5887748 3', mRNA sequence.
ACCESSION BM998354
VERSION   BM998354.1 GI:19723255
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 638)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
DNA Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5887748"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DT1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP DT1 is a normalized cDNA library containing the
following tissue(s): Metastatic Chondrosarcoma in Lung.
The library was constructed according to Bonaldo, Lennon

```

and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTV3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCCG.

TAG LIB=UI-H-DT1
TAG_TISSUE=lung metastatic chondrosarcoma
TAG_SEQ=AACTGTTCCG"

BASE COUNT 108 a 185 c 201 g 144 t
ORIGIN

Query Match		83.4%;	Score 617.2;	DB 12;	Length 638;
Best Local Similarity		99.5%;	Pred. NO. 1.3e-128;		
Matches 619;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	119	GTTCAGCTGAAGCTGTCGGTGGCGCTACGGGCGCATCCCGGGG	178		
DB	638	GTCAAGCTGAAGCTGTCGGTGGCGCTACGGGCGCATCCCGGGG	579		
QY	179	CGCGTGTCTTCATCGACGCTTGACCTACCGACAAGCTGGTCAGCTTCTACTGGA	238		
DB	578	CGCGTGTCTTCATCGACGCTTGACCTACCGACAAGCTGGTCAGCTTCTACTGGA	519		
QY	239	GACCTACGGGCGCGAGCTACCGCTAACGTGTCGGCGCATGGGCTGCAGGAGTGGC	298		
DB	518	GACCTACGGGCGCGAGCTACCGCTAACGTGTCGGCGCATGGGCTGCAGGAGTGGC	459		
QY	299	CGGGCAGCTGCAGGGCGGCACGACAGGGCTCTGGAGCGGCCAGCTGGGATCCAGGC	358		
DB	458	CTGGCAGCTGCAGGGCGGCACGACAGGGCTCTGGAGCGGCCAGCTGGGATCCAGGC	399		
QY	359	CCCTCCTCAGTCCGSCAGCAAGCCAGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCT	418		
DB	398	CCCTCCTCAGTCCGSCAGCAAGCCAGCCCTGCACTTTATAGACCAGCACCGGGCTGCGCT	339		
QY	419	TATCGGAGGCTCACAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCTCTGAC	478		
DB	338	TATCGGAGGCTCACAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCTCTGAC	279		
QY	479	GGATGAGCAGTACACGCGAGTGGGGCCGAGCCACCAACCAAGCAAGATGCGAAGCT	538		
DB	278	GGATGAGCAGTACACGCGAGTGGGGCCGAGCCACCAACCAAGCAAGATGCGAAGCT	219		
QY	539	CTTCAGTTTACACACGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAG	598		
DB	218	CTTCAGTTTACACACGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAG	159		
QY	599	GGAGTCCCAGTCTTACTGTTGAGGACCTGGAGCGGAGCTGAGGCTCTTCCCAGCAAC	658		
DB	158	GGAGTCCCAGTCTTACTGTTGAGGACCTGGAGCGGAGCTGAGGCTCTTCCCAGCAAC	99		
QY	659	ACTCCGGTCCAGCCCTGGCAATCCCAACCAATCATCTGAATCTTTTATACACA	718		
DB	98	ACTCCGGTCCAGCCCTGGCAATCCCAACCAATCATCTGAATCTTTTATACACA	39		
QY	719	ATATACGAAAAGCCAGCTTGA	740		
DB	38	ATATACGAAAAGCCAGCTTGA	17		

Search completed: January 29, 2004, 13:05:55
Job time : 1480.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:02:31 ; Search time 39.6087 Seconds
(without alignments)
8246.264 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtcggctgcagcgggg.....atcgaaagcagcttgaa 740

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	100.0	740	4	US-09-340-620A-48
2	585	79.1	585	4	US-09-340-620A-50
3	336	45.4	777	4	US-09-340-620A-60
4	316.2	42.7	579	4	US-09-340-620A-62
5	53	7.2	909	4	US-09-252-991A-2958
6	53	7.2	1218	4	US-09-252-991A-3146
7	53	7.2	1380	4	US-09-252-991A-2822
8	53	7.2	1500	4	US-09-252-991A-3050
9	50.6	6.8	2079	4	US-09-252-991A-3097
10	49.8	6.7	33529	3	US-09-144-085-3
11	49.2	6.6	50937	3	US-09-428-517-1
12	49	6.6	2610	4	US-09-252-991A-1107
13	49	6.6	3114	4	US-09-252-991A-977
14	49	6.6	3195	4	US-09-252-991A-937
15	48.8	6.6	468	4	US-09-252-991A-10961
16	48.8	6.6	795	4	US-09-252-991A-11104
17	48.8	6.6	1044	4	US-09-252-991A-11159
18	48.8	6.6	1557	4	US-09-252-991A-11093
19	48	6.5	30001	1	US-08-125-468-1
20	48	6.5	30001	2	US-08-474-933-1
21	47.8	6.5	558	4	US-09-252-991A-6608
22	47.8	6.5	645	4	US-09-252-991A-6723
23	47.8	6.5	984	4	US-09-252-991A-6664
24	46.6	6.3	800	5	PCT-US95-04801-4
25	46.2	6.2	23673	4	US-09-773-816-1
26	45	6.1	876	4	US-09-252-991A-12184
27	45	6.1	1392	4	US-09-252-991A-12249

c	28	45	6.1	1407	4	US-09-252-991A-12361	Sequence 12361, A
	29	44.8	6.1	780	4	US-09-252-991A-3741	Sequence 3741, Ap
c	30	44.8	6.1	960	4	US-09-252-991A-3730	Sequence 3730, Ap
	31	44.8	6.1	1089	4	US-09-252-991A-3785	Sequence 3785, Ap
c	32	44.6	6.0	1656	3	US-09-385-028-14	Sequence 14, Appl
	33	44.6	6.0	1656	4	US-09-726-614-14	Sequence 14, Appl
c	34	44.6	6.0	15079	3	US-09-385-048-1	Sequence 1, Appl
	35	44.6	6.0	15079	4	US-09-726-614-1	Sequence 1, Appl
c	36	44.4	6.0	1512	3	US-08-911-853-6	Sequence 6, Appl
	37	44.4	6.0	1512	3	US-09-479-409-6	Sequence 6, Appl
c	38	44.4	6.0	1512	4	US-09-479-453-6	Sequence 6, Appl
	39	44.4	6.0	17612	3	US-08-911-853-29	Sequence 29, Appl
	40	44.4	6.0	17612	3	US-09-479-409-29	Sequence 29, Appl
	41	44.4	6.0	17612	4	US-09-479-453-29	Sequence 29, Appl
c	42	44.2	6.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
	43	44	5.9	921	4	US-09-252-991A-13020	Sequence 13020, A
	44	44	5.9	973	4	US-09-482-273-13	Sequence 13, Appl
c	45	44	5.9	1491	4	US-09-252-991A-1345	Sequence 1345, Ap

ALIGNMENTS

RESULT 1
US-09-340-620A-48
; Sequence 48, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-340-620A-48

Query Match 100.0%; Score 740; DB 4; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.4e-157; Indels 0; Gaps 0;
Matches 740; Conservative 0; Mismatches 0

QY	1	CGCGTCCGCTCGAGCGGGTTCAGCGCGGCGAGCGCGCGGGATCTCTGAGGCATCGGCG 60
Db	1	CGCGTCCGCTCGAGCGGGTTCAGCGCGGCGAGCGCGCGGGATCTCTGAGGCATCGGCG 60
QY	61	GGCGCGCGACGCCATCTCTGGATGGCGTGGAGAACTGACCGCGAGGAGCTCAAGAAGT 120
Db	61	GGCGCGCGACGCCATCTCTGGATGGCGTGGAGAACTGACCGCGAGGAGCTCAAGAAGT 120
QY	121	TCAAGCTGAAGTCTGTCTGGTCCGCTCGCGAGGGCTACCGCGGCATCCCGCGGGCG 180
Db	121	TCAAGCTGAAGTCTGTCTGGTCCGCTCGCGAGGGCTACCGCGGCATCCCGCGGGCG 180
QY	181	CGCTGCTGCCATGAGCGCTTGGACCTCACCGACAAGCTGTTCAGCTTCTTACCTGAGA 240
Db	181	CGCTGCTGCCATGAGCGCTTGGACCTCACCGACAAGCTGTTCAGCTTCTTACCTGAGA 240
QY	241	CCTACGCGCGAGCTCACCGCTAACTGCTGCGGACATGGGCTTCAGGAGATGGCGG 300

[illegible]

```

RESULT 2
US-09-340-620A-50
; Sequence 50, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340, 620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-50

```

	Query Match	79.1%;	Score 585;	DB 4;	Length 585;
	Best Local Similarity	100.0%;	Pred. No. 8.5e-123;		
	Matches 585;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	54	ATGGGGCGCGCGGACGCCATCCTGGATGCGGTGGAGAACTGACCCCGGAGGATC	113		
Db	1	ATGGGGCGCGCGGACGCCATCCTGGATGCGGTGGAGAACTGACCCCGGAGGATC	60		
QY	114	AAGAAGTTCAAGCTGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGGCATCCCG	173		
Db	61	AAGAAGTTCAAGCTGAAGCTGCTGTCGGTCCGCTGCGGAGGGCTACGGGGCATCCCG	120		

QY	174	CGGGGCGCTGTGTCTCCATGAGCGCTTGGACCTCAACGACAAGCTGGTCAGCTTCTAC	233
Db	121	CGGGGCGCGTGTGTCCATGAGCGCTTGGACCTCAACGACAAGCTGGTCAGCTTCTAC	180
QY	234	CTGAGACCTTACGGGCGCGAGCTCACCGCTAAAGTGTGCGGACATGGGGCTCGAGAG	293
Db	181	CTGAGACCTTACGGGCGCGAGCTCACCGCTAAAGTGTGCGGACATGGGGCTCGAGAG	240
QY	294	ATGGCGGGGAGCTGCAGGCGGCGACACACACGAGGCTCTGGAGCGGGCCAGCTGGGATC	353
Db	241	ATGGCGGGGAGCTGCAGGCGGCGACACACGAGGCTCTGGAGCGGGCCAGCTGGGATC	300
QY	354	CAGGCCCTCCTCAGTCGGCAGCCAAAGCCAGGCTGCACCTTTATAGACCAGCACGGGCT	413
Db	301	CAGGCCCTCCTCAGTCGGCAGCCAAAGCCAGGCTGCACCTTTATAGACCAGCACGGGCT	360
QY	414	GCCTTATCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGATC	473
Db	361	GCCTTATCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGATC	420
QY	474	CTGACGGATCAGCAGTACCAAGCAGTGCGGGCGAGCCACCAACCAAGCAAGATCGG	533
Db	421	CTGACGGATGAGCAGTACCAAGCAGTGCGGGCGAGCCCAACCAAGCAAGATCGG	480
QY	534	AAGCTCTTCAGTTTCAACCAAGCCTGGAACTGGAACCTGCAAGGACTTGCTCCTCGAGGCC	593
Db	481	AAGCTCTTCAGTTTCAACCAAGCCTGGAACTGGAACCTGCAAGGACTTGCTCCTCGAGGCC	540
QY	594	CTAAGGGAGTCCAGTCTCACTACCTGGGAGCCTGAGCGGGAGC	638
Db	541	CTAAGGGAGTCCAGTCTCACTACCTGGGAGCCTGAGCGGGAGC	585

```

RESULT 3
US-09-340-620A-60
/ Sequence 60, Application US/09340620A
/ Patent No. 6482933
/ GENERAL INFORMATION:
/ APPLICANT: Bertin, John
/ TITLE OF INVENTION: NOVEL MOLECULES OF THE
/ FILE REFERENCE: 07334-124001
/ CURRENT APPLICATION NUMBER: US/09/340,620A
/ CURRENT FILING DATE: 1999-06-28
/ PRIOR APPLICATION NUMBER: US 09/245,281
/ PRIOR FILING DATE: 1999-02-05
/ PRIOR APPLICATION NUMBER: US 09/207,359
/ PRIOR FILING DATE: 1998-12-08
/ PRIOR APPLICATION NUMBER: US 09/099,041
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: US 09/019,942
/ PRIOR FILING DATE: 1998-02-06
/ NUMBER OF SEQ ID NOS: 71
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 60
/ LENGTH: 777
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (89)...(667)
US-09-340-620A-60

```

	Query Match	45.4%	Score 336;	DB 4;	Length 777;
	Best Local Similarity	72.94;	Pred. No. 7.4e-67;		
	Matches 461;	Conservative 0;	Mismatches 165;	Indels 6;	Gaps 2
QY	18	GGGTGAGCGCGCAGCGCGCGGGGATCTCTGAGCCATGGGGCGCGCGCGAGGCCATC	77		
DB	53	GAGTAAAGGTGACCGCGGTGCCACCCAGAGCCATGGGGGGGACGAGATGCCATC	112		
QY	78	CTGGATGCGCTGAGAAACCTGACGGCGGAGGCTCAAGAAGTTCAAGCTGAAGCTGCTG	137		

Db 113 CTGGAGCGCTCTTGAACCTTGTGAGGGGATGAACCTCAAAAAGTTCAAGATGAAGCTGCTG 172
Qy 138 TCGGTGCCCGCTCGCGAGGCTACGGCGGCATCCCGCGGGCGCGCTGTGTCTCCATGGAC 197
Db 173 ACAGTGAACCTGGAGAAGCTATGGCGCATCCACCGCGGGCCCTGCTGCAGATGGAC 232
Qy 198 GCGTTGGACCTCACCGACAAGCTGTGCTTACCTGGAGACCTTACCGCGCCGAGCTC 257
Db 233 GCCATAGATCTCACTGACAACTTGTGAGTCTATCTGAGTCTGATGCTTGGAGCTC 292
Qy 258 ACCGCTAAGTCTGTGCGGCACATGGCTGCGAGGAGATGGCGGCGAGCTGCGAGCGGCC 317
Db 293 ACAATGACTGTGTAGAGACATGGGCTTACAGGAGCTGGCTGAGCAGCTGCAAAAGCG--- 349
Qy 318 AGCACCAGGCTCTGGAGCGCGCGCAGCTGGGATCGGATCCAGCGCCCTCTCAGTCCGAGCGC 377
Db 350 ACTAAGAGAGCTGTGAGCTGTGGCAGCTGCGAGCCAGTGTCCCTGCTCAGAGTACAGCC 409
Qy 378 AAGCCAGGCTCGACTTTATAGACCAAGCACCGGGCTGCGCTTATCGCGAGGCTCAAAAC 437
Db 410 AGAACAGG---ACACTTTGTGACCAAGCACAGCAAGCACTCATTTGCCAGGGTCAAGAA 466
Qy 438 GTTGAGTGTCTGTGATCTGTACGGAAGTCTTACCGGATGAGCAGTACCAAGCA 497
Db 467 GTGGACGGAGTGTGATCTTGTGATGCGAGTGTGCTGACTGAAGGACAGTACCAAGCA 526
Qy 498 GTGGCGCGGAGCCACCAAGCAAGATGCGGAAGCTCTTCAAGTTTACACCGAGCC 557
Db 527 GTTGTGAGACACCAAGCAAGATGAGGAAGCTCTTCAAGCTTTGTTCATCC 586
Qy 558 TGGAACTGGACCTGCAAGGACTTGTCTTCCAGCGCCCTAAGGAGTCCAGTCTTACCTG 617
Db 587 TGGAACTGACCTGCAAGGACTCTCTCCAGGCTTGAAGAAATACATCCCTACTTG 646
Qy 618 GTGGAGGACTGAGCGGAGCTGAGCTCCTT 649
Db 647 GTGATGGACCTGGAGCAGAGCTGAGGTATCTT 678

RESULT 4
US-09-340-620A-62
; Sequence 62, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340, 620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-340-620A-62

Query Match 42.7%; Score 316.2; DB 4; Length 579;
Best Local Similarity 73.7%; Pred. No. 1.9e-62;
Matches 431; Conservative 0; Mismatches 148; Indels 6; Gaps 2;
Qy 54 ATGGGGGGCGCGGACGCGCATCTTGGATGCGCTGAGAACCTGACCGCGGAGGAGCTC 113
Db 1 ATGGGGGGCGGACGAGATGCCATCTCTGAGCGCTCTTGAAGAACTTGTGAGGGATGAATC 60

Qy 114 AAGAAAGTTCAAGCTGAAGCTGCTGTGCTGCGGTGCGGAGGCTACGGGGCGCATCCG 173
Db 61 AAAAAGTTCAAGATGAAGCTGTGACAGTGAACCTGCGAGAAGGCTATGGGGCGCATCCA 120
Qy 174 CGGGCGCGCTGCTGCTCCATGAGACGCTTGGACCTCACCGACAAGCTGTGAGCTTCTAC 233
Db 121 CCGGGGGCCCTGCTGCAGATGAGACGCCATAGATCTCACTGACAAACTTGTGAGCTACTAT 180
Qy 234 CTGGAGACCTTACGGCGCGAGCTCAACCGCTAACGCTGTGCGGACATGCGGCTGAGGAG 293
Db 181 CTGGAGTGTATGGCTTGGAGCTCAATGCTGTGCTTAGAGACATGGCTTACAGGAG 240
Qy 294 ATGGCGGGGAGCTGCGAGCGGCCACGACACAGGCTCTGAGAGCGCGCAGCTGGGATC 353
Db 241 CTGGCTGAGCAGCTGCAAAAG---ACTAAAGAGAGTCTGGAGCTGTGGCAGCTGCAGCC 297
Qy 354 CAGGCGCCCTCTCAGCTCGGCGAGCCAGCGCTGCTTATAGACAGACACCGGGCT 413
Db 298 AGTGTCCCTGCTCAGAGTACAGCCAGAACAGG---ACACTTTGTGAGCAGCAGCAAGCAA 354
Qy 414 GCGCTTATCGCGAGGCTCACAAACGTTGAGTGGCTGTGATGCTGTCTGTACGGGAAGTCT 473
Db 355 GCACCTCATTTGCAGGCTCACAGAGTGGACGGAGTGTGATGCTTGTGATGGCAGTGTG 414
Qy 474 CTGACGGATGAGCAGTACAGGCGAGTGGCGGCGGAGCCACCAACCAAGCAGATGCGG 533
Db 415 CTGACTGAAGGACAGTACAGGCGAGTTCGTGAGAGACCCAGCGCCAGCAAGATGAGG 474
Qy 534 AGCTCTTTCAGTTTACACAGGCTTGGAGCTGAGCTGCAAGGACTTGTCTCTCCAGGCG 593
Db 475 AAGCTCTTTCAGTTTGTTCATCTCTGGAACCTGAGCTGCAAGGACTTCTCTCTCCAGGCG 534
Qy 594 CTAAGGAGTCCAGCTCTTACCTGCTGGAGGACCTGGAGCGGAGC 638
Db 535 TTGAAGGAATAACATCCCTACTTGTGATGGACCTGGAGCAGC 579

RESULT 5
US-09-252-991A-2958
; Sequence 2958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107195.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2958
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2958

Query Match 7.2%; Score 53; DB 4; Length 909;
Best Local Similarity 49.1%; Pred. No. 0.0027;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGGAGCGCCATCTGATGCGTGGAGAACCTGACCGCGGAGGCTCAAGAAGTTC 122
Db 576 GCGCGCGCTCTCACCACCGAATACGAAGCGCGGCGGTCAAGGAAGTTCGCGCGCTTC 635
Qy 123 AAGCTGAAGCTGCTGTGCGTGTGCGGAGGCTACGGCGGATCCCGCGGCGCGCG 182
Db 636 CGCTCCTACTGCTGTTTGGAGCAGAGCTTCCAGGACTACGTACGTTCTCTCCAGGGCAAC 695
Qy 183 CTGCTGTCCATGGAGCGCTTGGACCTTCCAGCAAGAGCTGGTCAAGCTTCTTCTGGAGACC 242

Db 696 GACCGCTACCAAGCCCTCGACAGCGCGCCAAACCCCGAGCGCTTCATGAGAACTG 755
 QY 243 TAGCGCGCGAGCTCACCGCTACGCTGCTGCGGACATGGGCTTCAGAGATGCCGGG 302
 Db 756 CAGCGCGCGCGTACGCGCCAGCCGATCGCGAGTACGCGCGCGCAAGGTGGCGCAGATCGCCAGA 815
 QY 303 CAGCTGCGAGCGCGCACGACCGAGGCTCTGAGCGCGCGCCAGCT 347
 Db 816 CAGATGCAGACTACCAAGCGCGTGCAGCGCGCGCGCGCGCGCGCGCT 860

RESULT 6

US-09-252-991A-3146
 ; Sequence 3146, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 3146
 ; LENGTH: 1218
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3146

Query Match 7.2%; Score 53; DB 4; Length 1218;
 Best Local Similarity 49.1%; Pred. No. 0.0029;
 Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 63 GCGCGCGCGCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTTC 122
 Db 925 GCGCGCGCTCTACCAACCAATACGAGCGCGAGCGGTCAAGGAAGTCCGCGGTTC 984
 QY 123 AAGCTGAAGCTGTGCTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGCGCGCG 182
 Db 985 CGCTCTACTCTGCTGCTGAGCAGAGCTTCCAGACTAGTCACTTCTTCCAGGGCAAC 1044
 QY 183 CTGCTGTTCATGAGCGCTTGGACCTACCGAACAGCTGCTGAGCTTCTTACCTGAGACC 242
 Db 1045 GACCGCTACCAAGACGCGCTGGACAGCGCGCGCAACCCCGAGCGCTTCATGAGGAATG 1104
 QY 243 TAGCGCGCGAGCTCACCGCTACGCTGCTGCGGACATGGGCTTCAGAGATGCCGGG 302
 Db 1105 CAGCGCGCGGTACGCGCACCGCTACGCGAGTACGCGCGCGCGCGCGCGCGCT 1164
 QY 303 CAGCTGACGCGCGCGCACGACCGAGGCTCTGAGCGCGCGCCAGCT 347
 Db 1165 CAGATGCAGACTACCAAGCGCGTGCAGCGCGCGCGCGCGCGCGCT 1209

RESULT 7

US-09-252-991A-2822/c
 ; Sequence 2822, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 2822
 ; LENGTH: 1380
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-2822

Query Match 7.2%; Score 53; DB 4; Length 1380;
 Best Local Similarity 49.1%; Pred. No. 0.003;
 Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 63 GCGCGCGAGCGCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTTC 122
 Db 321 GCGCGCGCTCTCACCAACCAATACGAGCGCGCAAGCGGTCAAGGAAGTCCGCGGTTC 262
 QY 123 AAGCTGAAGCTGTGCTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGGCGCG 182
 Db 261 CGCTCTACTCTGCTGCTGAGCAGAGCTTCCAGCACTACGTCAGCTTCTCCAGGGCAAC 202
 QY 183 CTGCTGTCCATGGAGCGCTTGGACCTCACCAACAGCTGGTTCAGCTTCTTACTCTGAGACC 242
 Db 201 GACCGCTACAGAACCGCTTGGACAGCGCGCGCAACCCCGAGCGCTTCATGAGGAATG 142
 QY 243 TAGCGCGCGAGCTCACCGCTTACGCTGCTGCGCGACATGGGCTTCAGAGATGCCGGG 302
 Db 141 CAGCGCGCGGTACGCCACCGATCCGCGAGTACGCCCGCAAGGTGGCGCAGATGCCCAGA 82
 QY 303 CAGCTGAGCGCGCGCACGACCGAGGCTCTGAGCGCGCGCCAGCT 347
 Db 81 CAGATGCAGACTACCAAGCGCGTGCAGCGCGCGCGCGCGCGCT 37

RESULT 8

US-09-252-991A-3050
 ; Sequence 3050, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 3050
 ; LENGTH: 1500
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3050

Query Match 7.2%; Score 53; DB 4; Length 1500;
 Best Local Similarity 49.1%; Pred. No. 0.003;
 Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 63 GCGCGCGAGCGCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTTC 122
 Db 920 GCGCGCGCTCTCACCAACCAATACGAGCGCGCAAGCGGTCAAGGAAGTCCGCGGTTC 979
 QY 123 AAGCTGAAGCTGTGCTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGGCGCG 182
 Db 980 CGCTCTACTCTGCTGCTGAGCAGAGCTTCCAGCACTACGTCAGCTTCTTCCAGGGCAAC 1039
 QY 183 CTGCTGTCCATGGAGCGCTTGGACCTCACCAACAGCTGGTTCAGCTTCTTACTCTGAGACC 242
 Db 1040 GACCGCTACAGAACCGCTTGGACAGCGCGCGCGCAACCCCGAGCGCTTCATGAGGAATG 1099
 QY 243 TAGCGCGCGAGCTCACCGCTTACGCTGCTGCGGACATGGGCTTCAGAGATGCCGGG 302
 Db 1100 CAGCGCGCGGTACGCCACCGATCCGCGAGTACGCCCGCAAGGTGGCGCAGATGCCCAGA 1159

QY 326 GGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTCAGTCGGCAGCCCAAGCCAGG 385
 Db 1510 TCGGCCCCGTACCGCTGCTGGGTACGGGGCCGACTCTCAGCGCGCGCGCGCGCG 1451
 QY 386 CC 387
 Db 1450 CC 1449

Query Match 6.6%; Score 49; DB 4; Length 2610;
 Best Local Similarity 49.8%; Pred. No. 0.027;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

RESULT 12
 US-09-252-991A-1107/c
 ; Sequence 977, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 1107
 ; LENGTH: 2610
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-1107

QY 62 CGCGCGCGAGCGCATCTCGATGCGCTGGAGAACCTTGACCGCGGAGAGCTCAAGAAAGTT 121
 Db 1942 CGAGCGCGGATTCCGCTGCACGCGCCGCGGAGAGCTGAACCATCCGCTGTGGCTGGAGCG 1883
 QY 122 CAAGCTGAAGTCTGCTCGTGGTCCGCTGCGGAGGCTACGGCGCATCCCGGGGGCGC 181
 Db 1882 CATCGCCCACTGCGCGCGGAGCTTCCTGTCTCTTCTACTACCGCGCGCTGTCTCGGCGC 1823
 QY 182 GCTGCTGTCCATGAGCGCTTGAGACCTTGAGACCTCACGACAGCTGGTCTACTCTGAGAG 241
 Db 1822 CGAGCTGCTGCGCTGCGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1763
 QY 242 CTACGCGCGCGAGCTCACCGCTAACGCTGCTGCGCGCATGGGCGCTGCGAGAGATGCGCGG 301
 Db 1762 CTACGCGCGAGAGCGCGCGGAGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1703
 QY 302 GCAGCTGCA 310
 Db 1702 GACCTGCA 1694

Query Match 6.6%; Score 49; DB 4; Length 2610;
 Best Local Similarity 49.8%; Pred. No. 0.027;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

RESULT 13
 US-09-252-991A-977
 ; Sequence 977, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 977

QY 62 CGCGCGCGAGCGCATCTCGATGCGCTGGAGAACCTTGACCGCGGAGAGCTCAAGAAAGTT 121
 Db 1281 CGAGCGCGGATTCCGCTGCACGCGCCGCGGAGAGCTGAACCATCCGCTGTGGCTGGAGCG 1340
 QY 122 CAAGCTGAAGTCTGCTCGTGGTCCGCTGCGGAGGCTACGGCGCATCCCGCGGGGGCGC 181
 Db 1341 CATCGCCCACTGCGCGCGGAGCTTCCTGTCTCTTCTACTACCGCGCGCTGTCTCGGCGC 1400
 QY 182 GCTGCTGTCCATGAGCGCTTGAGACCTTGAGACCTCACGACAGCTGGTCTGCTGAGAG 241
 Db 1401 CGAGCTGCTGCGCTGCGCGCGGAGCTTCCTGTCTCTTCTACTACCGCGCGCTGTCTCGGCGC 1460
 QY 242 CTACGCGCGCGAGCTCACCGCTAACGCTGCTGCGGAGATGGGCGCTGCGAGGAGATGCGCGG 301
 Db 1461 CTACGCGCGAGAGCGCGCGGAGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1520
 QY 302 GCAGCTGCA 310
 Db 1521 GACCTGCA 1529

Query Match 6.6%; Score 49; DB 4; Length 3114;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

US-09-252-991A-977

RESULT 14
 US-09-252-991A-937
 ; Sequence 937, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 937
 ; LENGTH: 3195
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-937

QY 62 CGCGCGCGAGCGCATCTCGATGCGCTGGAGAACCTTGACCGCGGAGAGCTCAAGAAAGTT 121
 Db 478 CGAGCGCGGATTCCGCTGCACGCGCCGCGGAGAGCTGAACCATCCGCTGTGGCTGGAGCG 537
 QY 122 CAAGCTGAAGTCTGCTCGTGGTCCGCTGCGGAGGCTACGGCGCATCCCGCGGGGGCGC 181
 Db 538 CATCGCCCACTGCGCGCGGAGCTTCCTGTCTCTTCTACTACCGCGCGCTGTCTCGGCGC 597
 QY 182 GCTGCTGTCCATGAGCGCTTGAGACCTTGAGACCTCACGACAGCTGGTCTGCTGAGAG 241
 Db 598 CGAGCTGCTGCGCTGCGCGCGGAGCTTCCTGTCTCTTCTACTACCGCGCGCTGTCTCGGCGC 657
 QY 242 CTACGCGCGCGAGCTCACCGCTAACGCTGCTGCGGAGATGGGCGCTGCGAGGAGATGCGCGG 301
 Db 658 CTACGCGGAGAGCGCGCGGAGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
 QY 302 GCAGCTGCA 310

Query Match 6.6%; Score 49; DB 4; Length 3195;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Search completed: January 29, 2004, 13:11:29
Job time : 41.6087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:03:41 ; Search time 235.019 Seconds
(without alignments)
11475.528 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctcagcgggg.....atcacgaagcagcttgaa 740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_NA.*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	100.0	740	9	US-09-728-721-48
2	740	100.0	740	10	US-09-996-617-7
C 3	740	100.0	740	10	US-09-996-617-10
4	740	100.0	740	10	US-09-841-879B-4
C 5	740	100.0	740	10	US-09-841-879B-19
6	740	100.0	740	15	US-10-295-981-48
7	733	99.1	745	12	US-10-240-145-12
8	733	99.1	779	12	US-10-131-410-21
C 9	721.2	97.5	806	15	US-10-106-698-1144
10	697.6	94.3	811	9	US-09-925-301-278
11	585	79.1	585	9	US-09-728-721-50
12	585	79.1	585	10	US-09-996-617-9
13	585	79.1	585	10	US-09-841-879B-6
14	585	79.1	585	15	US-10-295-981-50
15	531	71.8	639	15	US-10-106-698-1145

16	384.6	52.0	432	9	US-09-925-299-648	Sequence 648, Appl
17	384.6	52.0	432	11	US-09-925-299-648	Sequence 648, Appl
18	336	45.4	777	9	US-09-728-721-60	Sequence 60, Appl
19	336	45.4	777	10	US-09-841-879B-1	Sequence 1, Appl
C 20	336	45.4	777	10	US-09-841-879B-18	Sequence 18, Appl
21	336	45.4	777	15	US-10-295-981-60	Sequence 60, Appl
22	316.2	42.7	579	9	US-09-728-721-62	Sequence 62, Appl
23	316.2	42.7	579	15	US-10-295-981-62	Sequence 62, Appl
C 24	309.8	41.9	321	10	US-09-880-107-1066	Sequence 1066, Ap
25	293.8	39.7	579	10	US-09-841-879B-3	Sequence 3, Appl
C 26	244.6	33.1	551	12	US-10-191-803-602	Sequence 602, App
27	169.4	22.9	270	11	US-09-965-621-27	Sequence 27, Appl
28	169.4	22.9	270	12	US-10-407-866-27	Sequence 27, Appl
29	151.2	20.4	437	12	US-10-062-674-985	Sequence 985, App
30	146.4	19.8	174	12	US-10-242-535A-45706	Sequence 45706, A
31	115	15.5	4200	10	US-09-388-221-3	Sequence 3, Appl
32	115	15.5	4287	10	US-09-996-617-5	Sequence 5, Appl
33	115	15.5	4287	10	US-09-931-071-5	Sequence 5, Appl
34	115	15.5	4332	10	US-09-388-221-5	Sequence 5, Appl
35	115	15.5	4422	10	US-09-388-221-1	Sequence 1, Appl
36	115	15.5	5100	11	US-09-956-712-12	Sequence 12, Appl
37	115	15.5	5444	10	US-09-996-617-1	Sequence 1, Appl
38	115	15.5	5444	10	US-09-931-071-1	Sequence 1, Appl
39	115	15.5	5444	11	US-09-956-712-3	Sequence 3, Appl
40	115	15.5	6531	11	US-09-956-712-11	Sequence 11, Appl
41	111.4	15.1	2657	11	US-09-895-298-22	Sequence 22, Appl
42	110.6	14.9	96649	11	US-09-956-712-10	Sequence 10, Appl
43	98.6	13.3	180	12	US-10-240-145-98	Sequence 98, Appl
C 44	74	10.0	394	9	US-09-864-761-21179	Sequence 21179, A
45	70.6	9.5	2708	12	US-10-407-866-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-09-728-721-48
; Sequence 48, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-728-721-48

Query Match 100.0%; Score 740; DB 9; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGCGTCCGCTCGAGCGGGTTCCTGGATGCGGCGGCGGCGGATCCTGGAGCCATGGGC	60
Db	1	CGCGTCCGCTCGAGCGGGTTCCTGGATGCGGCGGCGGCGGATCCTGGAGCCATGGGC	60
QY	61	GGCGCGCGACGCCATCCTGGATGCGTGGAGAACCTGACCGCGGAGGAGCTCAAGAAT	120


```

; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-10

```

Query Match	100.0%;	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0%;	Pred. No. 1.1e-196;		
Matches 740;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCGTCGGGCTGCAGCGGGGTGACGCGGCGAGCGCGCGGGATCTCTGGAGCCATGGGGC	60	
Db	740	CGCGTCGGGCTGCAGCGGGGTGACGCGGCGAGCGCGCGGGATCTCTGGAGCCATGGGGC	681	
Qy	61	GCGCGCGGACGCCATCTCTGGATCGCTGGAGAACTTGACCGCGGAGAGCTCAAGAAGT	120	
Db	680	GCGCGCGGACGCCATCTCTGGATCGCTGGAGAACTTGACCGCGGAGAGCTCAAGAAGT	621	
Qy	121	TCAAGCTGAAGCTCTGTCGTCGGCTGCGCGAGGGCTACGGGGGCATCCCGCGGGCGC	180	
Db	620	TCAAGCTGAAGCTCTGTCGTCGGCTGCGCGAGGGCTACGGGGGCATCCCGCGGGCGC	561	
Qy	181	CGCTGCTGTCCATGGAGCGCCTTGACCTTACCGACAAGCTGGTCAGCTTCTTACCTGGAGA	240	
Db	560	CGCTGCTGTCCATGGAGCGCCTTGACCTTACCGACAAGCTGGTCAGCTTCTTACCTGGAGA	501	
Qy	241	CTTACGGCGCGAGCTCACCGCTAACTGTCTGCGGACATGGGCTTGCAGGAGATGGCGC	300	
Db	500	CTTACGGCGCGAGCTCACCGCTAACTGTCTGCGGACATGGGCTTGCAGGAGATGGCGC	441	
Qy	301	GGCAGCTGCAGGCGGCGCACGCGACAGAGGCTCTGAGCGCGCGCAGCTGGGATCCAGGCC	360	
Db	440	GGCAGCTGCAGGCGGCGCACGCGACAGAGGCTCTGAGCGCGCGCAGCTGGGATCCAGGCC	381	
Qy	361	CTCTCTAGTCGGCAGCCAAAGCCAGGCTGCACTTTATAGACCAGCACGGGCTCGGCTTA	420	
Db	380	CTCTCTAGTCGGCAGCCAAAGCCAGGCTGCACTTTATAGACCAGCACGGGCTCGGCTTA	321	
Qy	421	TCGCGAGGGTCACAAACGTTTCAAGTGGCTGCTGGATGCTCTGTACGGGAGGTCCTGACGG	480	
Db	320	TCGCGAGGGTCACAAACGTTTCAAGTGGCTGCTGGATGCTCTGTACGGGAGGTCCTGACGG	261	
Qy	481	ATGAGCAGTACAGGCACTGCGGGCCGAGCCCAACACCCAAAGCAAGATGCGGAAGTCTCT	540	
Db	260	ATGAGCAGTACAGGCACTGCGGGCCGAGCCCAACACCCAAAGCAAGATGCGGAAGTCTCT	201	
Qy	541	TCAGTTTACACAGACCTGGAACTGGACCTGCAAGGACTTGCTCTCTCAGGCCCTTAAGGG	600	
Db	200	TCAGTTTACACAGACCTGGAACTGGACCTGCAAGGACTTGCTCTCTCAGGCCCTTAAGGG	141	
Qy	601	AGTCCCAGTCTACTGTTGGAGACCTGGAGCGAGCTGAGGCTCTCTTCCACGACACAC	660	
Db	140	AGTCCCAGTCTACTGTTGGAGACCTGGAGCGAGCTGAGGCTCTCTTCCACGACACAC	81	
Qy	661	TCCGGTCAGCCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAAT	720	
Db	80	TCCGGTCAGCCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAAT	21	
Qy	721	ATACGAAAGCCAGCTTGAA	740	
Db	20	ATACGAAAGCCAGCTTGAA	1	

RESULT 4
US-09-841-879B-4
; Sequence 4, Application US/09841879B
; Patent No. US20020142979A1

```

; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54) ... (638)
; US-09-841-879B-4

```

Query Match	100.0.0%;	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0.0%;	Pred. No. 1.1e-196;		
Matches 740;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCGTCGGGTGAGCGGGTGGCGCGAGCGCGCGGGATCTCTGGAGCCATGGGCG	60	
Db	1	CGCGTCGGGTGAGCGGGTGGCGCGAGCGCGCGGGATCTCTGGAGCCATGGGCG	60	
Qy	61	GGCGGCGGACGCCATCCTGGATGCGCTGGAGAACCTTGACCGCGGAGGAGCTCAAGAAGT	120	
Db	61	GGCGGCGGACGCCATCCTGGATGCGCTGGAGAACCTTGACCGCGGAGGAGCTCAAGAAGT	120	
Qy	121	TCAAGCTGAAGCTGCTGCTGGTCCGCTGCGGAGGGCTACGGCGCGATCCCGCGGGGCG	180	
Db	121	TCAAGCTGAAGCTGCTGCTGGTCCGCTGCGGAGGGCTACGGCGCGATCCCGCGGGGCG	180	
Qy	181	CGCTGCTGCTCCATGGAGCGCTTGAGCCTCACCGACAAGCTGGTCAGCTTCTACTCTGAGA	240	
Db	181	CGCTGCTGCTCCATGGAGCGCTTGAGCCTCACCGACAAGCTGGTCAGCTTCTACTCTGAGA	240	
Qy	241	CCTACGGCGCGAGCTCACCGCTAAACGTGTGCGCGACATGGGCTTCAGGAGATGGCCG	300	
Db	241	CCTACGGCGCGAGCTCACCGCTAAACGTGTGCGCGACATGGGCTTCAGGAGATGGCCG	300	
Qy	301	GGAGCTGAGGCGGCGACGCAACAGGGCTCTGGAGCGCGCCAGCTGGGATTCAGGCGCC	360	
Db	301	GGAGCTGAGGCGGCGACGCAACAGGGCTCTGGAGCGCGCCAGCTGGGATTCAGGCGCC	360	
Qy	361	CTCTCAGTCGGGAGCCAAAGCCAGGCGTGACATTTATAGACACGACCGGGCTCGGCTTA	420	
Db	361	CTCTCAGTCGGGAGCCAAAGCCAGGCGTGACATTTATAGACACGACCGGGCTCGGCTTA	420	
Qy	421	TCGCGAGGGTCAAAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCTCTGACGG	480	
Db	421	TCGCGAGGGTCAAAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCTCTGACGG	480	
Qy	481	ATGAGCAGTACAGGCGAGTCGGGCGCGAGCCACCAACCCAGCAAGATGCGGAAGCTCT	540	
Db	481	ATGAGCAGTACAGGCGAGTCGGGCGCGAGCCACCAACCCAGCAAGATGCGGAAGCTCT	540	
Qy	541	TCAGTTTTCACACAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGG	600	
Db	541	TCAGTTTTCACACAGCCTGGAACTGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGG	600	
Qy	601	AGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACAC	660	
Db	601	AGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACAC	660	
Qy	661	TCGGTTCAGGCCCTGGCAATCCCAACCAATCATCCTGAATCTGATCTTTTATACAAAT	720	
Db	661	TCGGTTCAGGCCCTGGCAATCCCAACCAATCATCCTGAATCTGATCTTTTATACAAAT	720	

QY 421 TCCGAGGCTCAAAACGTTGAGTGGCTGCTGATGCTCTGTACGGGAAGTCTCTGACGG 480
DB 421 TCCGAGGCTCAAAACGTTGAGTGGCTGCTGATGCTCTGTACGGGAAGTCTCTGACGG 480
QY 481 ATGACAGTACAGGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540
DB 481 ATGACAGTACAGGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540
QY 541 TCAGTTTACACAGCCTGGAACCTGACCTGCAAGGACTTGTCTCTCCAGGCGCTAAAGG 600
DB 541 TCAGTTTACACAGCCTGGAACCTGACCTGCAAGGACTTGTCTCTCCAGGCGCTAAAGG 600
QY 601 AGTCCAGTCTTACTGTGTGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
DB 601 AGTCCAGTCTTACTGTGTGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
QY 661 TCCGTCAGCCCTGCAATCCCAACCAATCATCTGAACTCTGATCTTTTATACAAAT 720
DB 661 TCCGTCAGCCCTGCAATCCCAACCAATCATCTGAACTCTGATCTTTTATACAAAT 720
QY 721 ATACGAAAGCCAGCTTGAA 740
DB 721 ATACGAAAGCCAGCTTGAA 740

RESULT 7
US-10-240-145-12
; Sequence 12, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 12
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(632)
US-10-240-145-12

Query Match 99.1%; Score 733; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 9.7e-195;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 GCGTCAGCGGGTGTAGCGGCGGAGCGGCGGAGTCTGTGGAGCATGGGCGCGCG 67
DB 2 GCGTCAGCGGGTGTAGCGGCGGAGCGGCGGAGTCTGTGGAGCATGGGCGCGCG 61
QY 68 CGACGCCATCTCTGATGCGCTGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGCT 127
DB 62 CGACGCCATCTCTGATGCGCTGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGCT 121
QY 128 GAAGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 187

DB 122 GAAGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
QY 188 GTCCATGAGCGCTCTGGACCTTACCGACAAGCTGTGTAGCTTTCTACCTGGAGACCTACGG 247
DB 182 GTCCATGAGCGCTCTGGACCTTACCGACAAGCTGTGTAGCTTTCTACCTGGAGACCTACGG 241
QY 248 CGCCAGCTCACCGCTAACGCTGCTGCGGACATGCGGCTGCGAGAGATGGCGGCGAGCT 307
DB 242 CGCCAGCTCACCGCTAACGCTGCTGCGGACATGCGGCTGCGAGAGATGGCGGCGAGCT 301
QY 308 CGAGCGCGCACGCAACAGGCTCTGGAGCGCGCGCAGCTGGGATCCAGGCGCTCTCTCA 367
DB 302 CGAGCGCGCACGCAACAGGCTCTGGAGCGCGCGCAGCTGGGATCCAGGCGCTCTCTCA 361
QY 368 GTCCGAGCCAGCGAGCGCTGCACTTTATAGACAGCACCGGGCTGGCTTTATCGCGAG 427
DB 362 GTCCGAGCCAGCGAGCGCTGCACTTTATAGACAGCACCGGGCTGGCTTTATCGCGAG 421
QY 428 GGTCAAAACGTTGAGTGGCTGCTGATGCTGTGTACGGGAAGTCTCTCAGTTT 547
DB 422 GGTCAAAACGTTGAGTGGCTGCTGATGCTGTGTACGGGAAGTCTCTCAGTTT 541
QY 488 GTACCAGCGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAAGTTT 547
DB 482 GTACCAGCGAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAAGTTT 541
QY 548 CACACAGCGCTGGAACCTGGAACCTGGAACCTGCTCTCTCCAGGCGCTTAAGGAGTCCCA 607
DB 542 CACACAGCGCTGGAACCTGGAACCTGGAACCTGCTCTCTCCAGGCGCTTAAGGAGTCCCA 601
QY 608 GTCTTACTGTGTGAGGAGCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCGGCTC 667
DB 602 GTCTTACTGTGTGAGGAGCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCGGCTC 661
QY 668 AGCCCTTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 727
DB 662 AGCCCTTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 721
QY 728 AAGCCAGCTTGAA 740
DB 722 AAGCCAGCTTGAA 734

RESULT 8
US-10-131-410-21
; Sequence 21, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-21
Query Match 99.1%; Score 733; DB 12; Length 779;
Best Local Similarity 100.0%; Pred. No. 9.8e-195;

```
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 CGGCTCAGCGGGTGTAGCGGGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 66
Db 47 CGGCTCAGCGGGTGTAGCGGGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 106
Qy 67 GCGACGCCATCTCTGGATGCGGTGGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGC 126
Db 107 GCGACGCCATCTCTGGATGCGGTGGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGC 166
Qy 127 TGAAGCTGTCTGCTGCGGTGCGGAGGAGCTTACCGGCGCATCCCGCGGGCGCGCTGC 186
Db 167 TGAAGCTGTCTGCTGCGGTGCGGAGGAGCTTACCGGCGCATCCCGCGGGCGCGCTGC 226
Qy 187 TGTCCATGAGCGCTTGGACCTCACCGAAGCTGTGCTACCTTGGAGAGCTTAGC 246
Db 227 TGTCCATGAGCGCTTGGACCTCACCGAAGCTGTGCTACCTTGGAGAGCTTAGC 286
Qy 247 GCGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTGAGGAGATGGCGGGCAGC 306
Db 287 GCGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTGAGGAGATGGCGGGCAGC 346
Qy 307 TGAGCGCGCAGCAGCAGCGGCTTGGAGCGCGCAGCTGGATGCCAGGCCCTTCCTC 366
Db 347 TGAGCGCGCAGCAGCAGCGGCTTGGAGCGCGCAGCTGGATGCCAGGCCCTTCCTC 406
Qy 367 AGTCGCGCGCAGCAGCGGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 426
Db 407 AGTCGCGCGCAGCAGCGGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 466
Qy 427 GGGTCACAAAGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 486
Db 467 GGGTCACAAAGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 526
Qy 487 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGCTTTCAGTT 546
Db 527 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGCTTTCAGTT 586
Qy 547 TCACACCGCTTGGAGTGGAGCTGCAAGGAGTGGCTTCCAGGCGCTTAAAGGAGTCC 606
Db 587 TCACACCGCTTGGAGTGGAGCTGCAAGGAGTGGCTTCCAGGCGCTTAAAGGAGTCC 646
Qy 607 AGTCTTACCTTGGTGGAGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 666
Db 647 AGTCTTACCTTGGTGGAGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 706
Qy 667 CAGCCCTCGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
Db 707 CAGCCCTCGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 766
Qy 727 AAAGCCAGCTTGA 739
Db 767 AAAGCCAGCTTGA 779
```

```
RESULT 9
US-10-106-698-1144/c
; Sequence 1144, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
```

```
; SEQ ID NO 1144
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-1144

Query Match 97.5%; Score 721.2; DB 15; Length 806;
Best Local Similarity 99.6%; Pred. No. 2e-191;
Matches 731; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 7 CGGCTCAGCGGGTGTAGCGGGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 66
Db 739 CGGCTCAGCGGGTGTAGCGGGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 680
Qy 67 GCGACGCCATCTCTGGATGCGGTGGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGC 126
Db 679 GCGACGCCATCTCTGGATGCGGTGGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGC 620
Qy 127 TGAAGCTGTCTGCTGCGGTGCGGAGGAGCTACGGCGCATCCCGCGGGCGCGCTGC 186
Db 619 TGAAGCTGTCTGCTGCGGTGCGGAGGAGCTACGGCGCATCCCGCGGGCGCGCTGC 560
Qy 187 TGTCCATGAGCGCTTGGACCTCACCGAAGCTGTGCTACCTTGGAGAGCTTAGC 246
Db 559 TGTCCATGAGCGCTTGGACCTCACCGAAGCTGTGCTACCTTGGAGAGCTTAGC 500
Qy 247 GCGCGGAGCTCACCGCTAACGCTGCTGCGGAGAGTGGCGGGCTGCGAGAGATGGCGGGCAGC 306
Db 499 GCGCGGAGCTCACCGCTAACGCTGCTGCGGAGAGTGGCGGGCTGCGAGAGATGGCGGGCAGC 440
Qy 307 TGCAGCGCGCAGCAGCAGCGGCTTGGAGCGCGCAGCTGGGATCCAGGCGCTTCTC 366
Db 439 TGCAGCGCGCAGCAGCAGCGGCTTGGAGCGCGCAGCTGGGATCCAGGCGCTTCTC 380
Qy 367 AGTCGCGCGCAGCAGCAGCGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 426
Db 379 AGTCGCGCGCAGCAGCAGCGCTTGGATGCTGTACCGGAGAGTGGCGGGCTGCGTTATCGCA 320
Qy 427 GGGTCACAAAGCTTGGATGCTGTACCGGAGAGTGGCGGGAGTGGCTTAAAGGAGTCC 486
Db 319 GGGTCACAAAGCTTGGATGCTGTACCGGAGAGTGGCTTAAAGGAGTGGCTTAAAGGAGTCC 260
Qy 487 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGCTTTCAGTT 546
Db 259 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGCTTTCAGTT 200
Qy 547 TCACACCGCTTGGAGTGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 606
Db 199 TCACACCGCTTGGAGTGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 140
Qy 607 AGTCCTTACCTTGGTGGAGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 666
Db 139 AGTCCTTACCTTGGTGGAGGAGCTTGGAGGAGTGGCTTCCAGCAACACTCCCGT 80
Qy 667 CAGCCCTCGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
Db 79 CA-CCCTTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 21
Qy 727 AAAGCCAGCTTGA 740
Db 20 AAAGCCAGCTTGA 7
```

```
RESULT 10
US-09-925-301-278
; Sequence 278, Application US/0925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
```

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 278
; LENGTH: 811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-278

Query Match 94.3%; Score 697.6; DB 9; Length 811;
Best Local Similarity 99.5%; Pred. No. 7.6e-185;
Matches 730; Conservative 1; Mismatches 0; Indels 3; Gaps 3;

QY	7	CGGCTGCACGGGTCAGCGCGGCGAGCGCGGGATCTCGAGCCATGGGCGCGCGC	66
DB	68	CGGCTGCACGGGTCAGCGCGGCGAGCGCGGGATCTCGAGCCATGGGCGCGCGC	127
QY	67	GGGAGCCATCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC	126
DB	128	GGAGCCATCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC	187
QY	127	TGAAGCTGCTGTCGTCGCGTCGCGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	186
DB	188	TGAAGCTGCTGTCGTCGCGTCGCGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGC	247
QY	187	TGTCCATGACGCTTGGACCTCACCGACAAGCTGTGAGCTTCTACCTGGAGACCTAGC	246
DB	248	TGTCCATGACGCTTGGACCTCACCGACAAGCTGTGAGCTTCTACCTGGAGACCTAGC	307
QY	247	GGCGCGAGCTCACCGCTAACGTCGTCGCGCATGCGGCTCGAGAGATGGCGCGGCGAGC	306
DB	308	GGCGCGAGCTCACCGCTAACGTCGTCGCGCATGCGGCTCGAGAGATGGCGCGGCGAGC	367
QY	307	TGACGCGCCACGACACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTC	366
DB	368	TGACGCGCCACGACACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCGCCCTCTC	427
QY	367	AGTCGCGACCCAGCGAGGCTGCACTTTATAGACGACCGGGCTGGCTTATCGGA	426
DB	428	AGTCGCGACCCAGCGAGGCTGCACTTTATAGACGACCGGGCTGGCTTATCGGA	487
QY	427	GGTCAACAAAGTTGAGTGGCTGCTGATGCTCTGTAACGGAGGTCCTGACGAGTACG	486
DB	488	GGTCAACAAAGTTGAGTGGCTGCTGATGCTCTGTAACGGAGGTCCTGACGAGTACG	547
QY	487	AGTACGAGGTCGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	546
DB	548	AGTACGAGGTCGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT	606
QY	547	TCACACGCTGGAACTGGAACCTGCAAGGATTTGCTCTCCAGGCGCTTAAGGAGTCCC	606
DB	607	TCACACGCTGGAACTGGAACCTGCAAGGATTTGCTCTCCAGGCGCTTAAGGAGTCCC	666
QY	607	AGTCTTACCTGTTGAGGAGCTGGAGCGAGCTGAGGCTCTCCAGCAACACTCCCGT	666
DB	667	AGTCTTACCTGTTGAGGAGCTGGAGCGAGCTGAGGCTCTCCAGCAACACTCCCGT	725
QY	667	CAGGCGCTGGCAATCCACCAATCATCTGATATCTGATCTTTTATACAAATATACGA	726
DB	726	CA-SCCTGGCAATCCACCAATCATCTGATATCTGATCTTTTATACAAATATACGA	784
QY	727	AAAGCAGCTTGAA	740
DB	785	AAAGCAGCTTGAA	798

RESULT 11
US-09-728-721-50
; Sequence 50, Application US/09728721
; Patent No. US20020061845A1

; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-728-721-50

Query Match 79.1%; Score 585; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	54	ATGGGCGCGCGCGCGAGCCATCTCGATGCGCTGGAGAACTCGACCGCGGAGGATC	113
DB	1	ATGGGCGCGCGCGCGAGCCATCTCGATGCGCTGGAGAACTCGACCGCGGAGGATC	60
QY	114	AAGAAAGTTCAAGCTGAAGCTGCTGTCGTCGCTGCGCGAGGCTACGGCGCATCCCG	173
DB	61	AAGAAAGTTCAAGCTGAAGCTGCTGTCGTCGCTGCGCGAGGCTACGGCGCATCCCG	120
QY	174	CGGGCGCGCTGCTGTCATGAGCGCTTGACCTCAGCGACCAAGCTGTCAGCTTCTAC	233
DB	121	CGGGCGCGCTGCTGTCATGAGCGCTTGACCTCAGCGACCAAGCTGTCAGCTTCTAC	180
QY	234	CTGGAGAGCTACCGCGCGAGCTCACCGCTAACGTCGTCGCGACATGGGCTCGAGGAG	293
DB	181	CTGGAGAGCTACCGCGCGAGCTCACCGCTAACGTCGTCGCGACATGGGCTCGAGGAG	240
QY	294	ATGGCGCGGAGCTGACGGCGGCGCACGACCAAGGCTCTGGAGCGCGCGAGCTGGATC	353
DB	241	ATGGCGCGGAGCTGACGGCGGCGCACGACCAAGGCTCTGGAGCGCGCGAGCTGGATC	300
QY	354	CAGGCGCTCTCAGTCGCGAGCGCAAGCGCTGCACTTTATAGACCAAGCGGCT	413
DB	301	CAGGCGCTCTCAGTCGCGAGCGCAAGCGCTGCACTTTATAGACCAAGCGGCT	360
QY	414	GCGCTTATCGGAGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTACGGGAAGTTC	473
DB	361	GCGCTTATCGGAGGTCACAAACGTTGAGTGGCTGCTGGATGCTGTACGGGAAGTTC	420
QY	474	CTGACGGATGAGCAGTACAGCAGTGGCGGCGGAGCCACCAAGCAAGATGGCG	533
DB	421	CTGACGGATGAGCAGTACAGCAGTGGCGGCGGAGCCACCAAGCAAGATGGCG	480
QY	534	AAGCTCTTCAGTTTACACCGAGCTGGAACTGGAGCTGCAAGGACTTGTCTCCAGGCC	593
DB	481	AAGCTCTTCAGTTTACACCGAGCTGGAACTGGAGCTGCAAGGACTTGTCTCCAGGCC	540
QY	594	CTAAGGGAGTCCAGTCTTACCTGTTGGAGGAGCTGGAGCGGAGC	638
DB	541	CTAAGGGAGTCCAGTCTTACCTGTTGGAGGAGCTGGAGCGGAGC	585

RESULT 12
US-09-996-617-9
; Sequence 9, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John

; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 585
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-981-50

Query Match 79.1%; Score 585; DB 15; Length 585;
 Best Local Similarity 100.0%; Pred. No. 2e-153;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	54	ATGGGGCGCGCGCGACCCATCTCGGTGAGTGGCTGAGAACCTGACCGCCGAGGATC 113
DB	1	ATGGGGCGCGCGCGACCCATCTCGGTGAGTGGCTGAGAACCTGACCGCCGAGGATC 60
QY	114	AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCTGCGGAGGGCTACCGGGCGCATCCCG 173
DB	61	AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCTGCGGAGGGCTACCGGGCGCATCCCG 120
QY	174	CGGGCGCGCTGCTGTCATGAGCGCTTGGACCTCACCGACAAAGCTGTGCTGACCTTCTAC 233
DB	121	CGGGCGCGCTGCTGTCATGAGCGCTTGGACCTCACCGACAAAGCTGTGCTGACCTTCTAC 180
QY	234	CTGGAGACCTACCGCGCGAGCTACCGCTTAAGCTGCTGGCGACATGAGGCTGCGAGGAG 293
DB	181	CTGGAGACCTACCGCGCGAGCTACCGCTTAAGCTGCTGGCGACATGAGGCTGCGAGGAG 240
QY	294	ATGGCGGGCAGCTGCGAGCGCGCCAGCCAGCGAGGCTCTGGAGCGCGCAGCTGGGATC 353
DB	241	ATGGCGGGCAGCTGCGAGCGCGCCAGCCAGCGAGGCTCTGGAGCGCGCAGCTGGGATC 300
QY	354	CAGGCCCTCTCCTCAGTCGGCAGCCAGCGAGGCTGCTGATGAGTGGCTGCTGATGAGTGG 413
DB	301	CAGGCCCTCTCCTCAGTCGGCAGCCAGCGAGGCTGCTGATGAGTGGCTGCTGATGAGTGG 360
QY	414	GGCTTTATCGCGAGGCTCAAAAGCTGAGTGGCTGCTGATGAGTGGCTGCTGATGAGTGG 473
DB	361	GGCTTTATCGCGAGGCTCAAAAGCTGAGTGGCTGCTGATGAGTGGCTGCTGATGAGTGG 420
QY	474	CTGACCGATGAGCAGTACAGGCGAGTGGCGGCGAGCGCCAGCCAGCCAGCAGGATGCGG 533
DB	421	CTGACCGATGAGCAGTACAGGCGAGTGGCGGCGAGCGCCAGCCAGCCAGCAGGATGCGG 480
QY	534	AAGCTCTTCAGTTTCAACAGCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCG 593
DB	481	AAGCTCTTCAGTTTCAACAGCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCG 540
QY	594	CTAAGGAGTCCAGTCTTACCTGAGTGGTGGAGGACCTGGAGCGGAGCCTGGAGCGGAGC 638
DB	541	CTAAGGAGTCCAGTCTTACCTGAGTGGTGGAGGACCTGGAGCGGAGCCTGGAGCGGAGC 585

RESULT 15
 US-10-106-698-1145
 ; Sequence 1145, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005PI
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 1145
 ; LENGTH: 639
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-106-698-1145

Query Match 71.8%; Score 531; DB 15; Length 639;
 Best Local Similarity 94.9%; Pred. No. 2.4e-138;
 Matches 595; Conservative 0; Mismatches 20; Indels 12; Gaps 4;

QY	124	AGCTGAAGCTGCTGTCGGTGGCTGCGCGAGGGCTACGGGCGCATCCCGCG-----G 176
DB	2	AGCTGAAGCTGCTGTCGGTGGCTGCGCGAGGGCTACGGGCGCATCCCGCG-----G 61
QY	177	GGCGCGCTGCTGCTCAGTGGACCGCTTGGACCTCACCGA---CAAGCTGGTGGCTTCTAC 233
DB	62	GATGCGCTGGAGAACTGACCGCGAGGAGCTCAAGAAAGTTCAAGCTGGTGGCTTCTAC 121
QY	234	CTGGAGACCTACCGCGCGAGCTCACCGCTAACGTGCTGCGGACATGAGGCTGCGAGGAG 293
DB	122	CTGGAGACCTACCGCGCGAGCTCACCGCTAACGTGCTGCGGACATGAGGCTGCGAGGAG 181
QY	294	ATGGCGGGCAGCTGCGAGCGCGCCAGCCAGCGAGGCTCTGGAGCGCGCAGCTGGGATC 353
DB	182	ATGGCGGGCAGCTGCGAGCGCGCCAGCCAGCGAGGCTCTGGAGCGCGCAGCTGGGATC 241
QY	354	CAGGCCCTCTCCTCAGTCGGCAGCGCAAGCAGCGCTGCACTTTATAGACCAGCACCGGGCT 413
DB	242	CAGGCCCTCTCCTCAGTCGGCAGCGCAAGCAGCGCTGCACTTTATAGACCAGCACCGGGCT 301
QY	414	GGCTTTATCGCGAGGCTCAAAAGCTGAGTGGCTGCTGATGAGTGGCTGCTGATGAGTGG 473
DB	302	GGCTTTATCGCGAGGCTCAAAAGCTGAGTGGCTGCTGATGAGTGGCTGCTGATGAGTGG 361
QY	474	CTGACCGATGAGCAGTACCGAGCAGTGGCGGCGAGCGCCAGCCAGCCAGCAGGATGCGG 533
DB	362	CTGACCGATGAGCAGTACCGAGCAGTGGCGGCGAGCGCCAGCCAGCCAGCAGGATGCGG 420
QY	534	AAGCTCTTCAGTTTCAACAGCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCG 593
DB	421	AAGCTCTTCAGTTTCAACAGCCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCG 480
QY	594	CTAAGGAGTCCAGTCTTACCTGAGTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCCA 653
DB	481	CTAAGGAGTCCAGTCTTACCTGAGTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCCA 540
QY	654	GCAACACTCCGGTCCAGCCCTGGCAATCCCAATCATCTGATCTGATCTTTTAT 713
DB	541	GCAACACTCCGGTCCAGCCCTGGCAATCCCAATCATCTGATCTGATCTTTTAT 599
QY	714	ACACAATATACGAAAGCAGCTTGAA 740
DB	600	ACACAATATACGAAAGCAGCTTGAA 626

Search completed: January 29, 2004, 13:44:37
 Job time : 236.019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 9.48584 Seconds
(without alignments)
3262.936 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARDAILDALENLTABEL.....LLQLALRESQSYLVEDLERS 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	20 AA148553	Human breast tumou
2	990	100.0	195	22 AA068525	Human novel cytol
3	990	100.0	195	22 AA000588	Human target of me
4	990	100.0	195	22 AAB20085	Human CARD-5 prote
5	990	100.0	195	23 AAU99353	Human caspase recr
6	990	100.0	195	23 AA017854	Human domain conta
7	990	100.0	195	24 ABG71635	Human caspase recr
8	990	100.0	205	22 AAG74647	Human colon cancer
9	873.5	88.2	176	22 AAR00594	Alternatively spli

10	759	76.7	190	21	AA043675	Human cancer assoc
11	684	69.1	193	22	AA000592	Mouse target of me
12	684	69.1	193	22	AA020086	Mouse CARD-5 prote
13	684	69.1	193	23	AAU99352	Mouse caspase recr
14	576	58.2	171	22	AAE00593	Rat target of meth
15	564.5	57.0	158	22	AAG74648	Human colon cancer
16	545	55.1	110	22	AAE00595	Alternatively spli
17	509	51.4	136	21	AA033881	Human colon cancer
18	448	45.3	91	22	AAE00589	Human target of me
19	439	44.3	84	22	AAE00591	Human target of me
20	305	30.8	76	23	AA021934	PAAD domain-contai
21	267	27.0	89	22	AA036608	Human FLEXHT-30 pr
22	267	27.0	89	23	AA017852	Pyrim domain conta
23	253	25.6	1397	22	AA072670	Human NB-ARC and C
24	250.5	25.3	1429	22	AA062571	Human CARD-7 polyp
25	250.5	25.3	1429	23	ABG97969	Leucine-rich repea
26	250.5	25.3	1429	23	ABG78472	Human leucine rich
27	250.5	25.3	1429	23	AB077916	Human leucine-rich
28	250.5	25.3	1429	24	ABG71631	Human caspase recr
29	250.5	25.3	1429	24	ABG71633	Human caspase recr
30	250	25.3	1442	22	AA072671	Human NB-ARC and C
31	250	25.3	1473	22	AA072669	Human NB-ARC and C
32	250	25.3	1473	22	AA072711	Human NAC beta iso
33	249.5	25.2	1429	23	ABG78455	Human caspase recr
34	249.5	25.2	1429	23	AA017855	Pyrim domain conta
35	248.5	25.1	1473	22	AA06758	Human G-protein co
36	240	24.2	442	21	AA024513	Human secreted pro
37	182	18.4	76	23	AA021935	PAAD domain-contai
38	153	15.5	65	21	AA024519	Human secreted pro
39	147	14.8	77	22	AA072516	Human PYRIN-1 doma
40	137	13.8	77	22	AA072515	Human PYRIN-1 doma
41	125.5	12.7	891	22	AA04546	Human G-protein co
42	125.5	12.7	1851	22	AAU00023	Human activated T-
43	125.5	12.7	1851	23	AA017860	Pyrim domain conta
44	122.5	12.4	781	20	AA03001	FMF associated pro
45	122.5	12.4	781	23	AA017853	Pyrim domain conta

ALIGNMENTS

RESULT 1
AA048553
ID AA048553 standard; Protein; 195 AA.
XX
AC AA048553;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human breast tumour-associated protein 14.
XX
KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;
KW treatment; tumour; cytostatic; medicament.
XX
OS Homo sapiens.
XX
PN DE19813839-A1.
XX
PD 23-SEP-1999.
XX
PF 20-MAR-1998; 98DE-1013839.
XX
PR 20-MAR-1998; 98DE-1013839.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;
XX
DR WPI; 1999-528981/45.
XX
PT N-PSDB; AA233631.
XX
PT Human nucleic acid sequences and protein products from tumor breast
tissue, useful for breast cancer therapy -

AAE00588;
02-JUL-2001 (first entry)
Human target of methylation-induced silencing-1 (TMS1) protein.
Human; target of methylation-induced silencing-1; TMS1; cytostatic;
anti-proliferative; apoptosis inducer; gene therapy; CpG island;
caspase-recruiting domain; CARD; cancer; breast.
Homo sapiens.
WO200129235-A2.
26-APR-2001.
18-OCT-2000; 2000WO-US28747.
18-OCT-1999; 99US-0159975.
(UYEM-) UNIV EMORY.
Vertino PM;
WPI; 2001-290922/30.
N-PSDB; AAD03889, AAD03890.
Novel gene TMS1, transcriptionally silenced due to increased
methylation useful for identifying subject at risk of developing tumor
characterized by abnormal methylation, for treating cancer by inducing
apoptosis -
Claim 85; Page 114; 124pp; English.
The invention relates to identification of target of methylation-induced
silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
abnormal methylation of a CpG island in its 5' regulatory region. TMS1
consists of a carboxy terminal caspase-recruiting domain (CARD) and
plays a role in induction of apoptosis. TMS1 gene and protein are useful
as tools for diagnosing and treating a subject at risk of developing
cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
abnormally low levels of TMS1 expression products. Unique fragments of
TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
TMS1 molecule is also useful for treating abnormal cell proliferation by
increasing TMS1 polypeptide level to an above normal level. The CpG
island region of TMS1 or its fragments are used to study the methylation
patterns apart from any coding region contained in it.
The present sequence is human target of methylation-induced silencing-1
(TMS1) protein.
Sequence 195 AA;
Query Match 100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAILDALENLTAEELKKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSY 60
QY 61 LETYGAEALTANVLRLDMGLQEMAGLQQAATHQSGGAAPAGIQAPPSAAKPGHLFIQDQRA 120
DB 61 LETYGAEALTANVLRLDMGLQEMAGLQQAATHQSGGAAPAGIQAPPSAAKPGHLFIQDQRA 120
QY 121 ALIARTVNVWLLDLYLGVKLTDEQVQAVRAEPTNPSKMKLFSFTPAWNWTKDILLQA 180
DB 121 ALIARTVNVWLLDLYLGVKLTDEQVQAVRAEPTNPSKMKLFSFTPAWNWTKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 4
AAB20085
ID AAB20085 standard; Protein; 195 AA.
XX
AC AAB20085;
XX
DT 23-APR-2001 (first entry)
XX
DE Human CARD-5 protein.
XX
KW CARD-5; caspase recruitment domain; human; cancer; infection;
KW autoimmune disease; neurological disease; haematological disease;
KW immune disease; inflammation; antitumour; antiseptic;
KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 111..181
FT /note= "CARD"
XX
PN WO200100826-A2.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17691.
XX
PR 28-JUN-1999; 99US-0340620.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-061973/07.
DR N-PSDB; AAF30007.
XX
PT Isolated intracellular proteins predicted to be involved in regulating
PT caspase activation are used for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PT haematological disorders -
XX
PS Claim 9; Fig 21; 208pp; English.
XX
CC The present sequence is that of human caspase recruitment domain 5
CC (CARD-5), an intracellular protein predicted to be involved in
CC regulating caspase activation. The sequence is predicted from an
CC isolated cDNA clone (see AAF30007). Methods of diagnosing and
CC treating patients suffering from a disorder associated with an
CC abnormal level or rate of apoptotic cell death, abnormal activity
CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
CC necrosis factor receptor complex or abnormal activity of a caspase
CC involve administering a compound that modulates the expression or
CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
CC therapy methods. Such disorders include cancer, viral infection,
CC autoimmune disorders, neurological diseases, haematological
CC disorders, inflammatory disorders and immune disorders. CARD-3,
CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
CC cell survival and cell growth. They can also be used to screen
CC drugs or compounds that modulate their activity or expression and
CC to treat disorders associated with insufficient or excessive
CC production of CARD-3, -4, -5 or -6 protein, or production of an
CC aberrant protein.
XX
SQ Sequence 195 AA;
Query Match 100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAILDALENLTAEELKKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSY 60

QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGTQAPPOSAAPKPLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGTQAPPOSAAPKPLHFDQHRA 120
 QY 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 DB 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 5
 AAU99353
 ID AAU99353 standard; Protein; 195 AA.
 AC AAU99353;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human caspase recruitment domain-5 (CARD-5) protein.
 KW Human; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX
 OS Homo sapiens.
 XX
 XX
 XX
 .FH Key Location/Qualifiers
 FT Domain 111..195
 FT /label= CARD_domain
 XX
 PN WO200244354-A2.
 XX
 XX
 PD 06-JUN-2002.
 XX
 XX 29-NOV-2001; 2001WO-US44894.
 PF
 PR 01-DEC-2000; 2000US-0728721.
 PR 24-APR-2001; 2001US-0841879.
 XX
 XX (MILL-) MILLENIUM PHARM INC.
 PA
 XX Bertin J;
 PI
 XX
 XX WPI; 2002-557538/59.
 DR N-PSDB; ABR87966.
 XX
 XX Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis -
 XX
 XX Claim 22; Fig 3; 100pp; English.
 XX
 CC The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteinyl aspartate-specific proteinases) are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CARDS and different
 CC subtypes of CARDS may confer binding specificity. CARD-5 is an

CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B
 CC (NF-kappaB) transcription factor pathway and binds the CARDS of
 CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
 CC activity and NF-kappaB activation, regulate cell growth and cell death
 CC and be used in gene therapy. The CARD-5 polypeptides are useful for
 CC identifying compounds which bind to them and modulate their activity and
 CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
 CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
 CC can be used to treat immune disorders such as chronic inflammatory
 CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
 CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
 CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
 CC (including tuberculosis and lepromatous leprosy) and in screening and
 CC detection assays. Modulators of CARD-5 activity or expression are also
 CC useful for treating autoimmune disorders, such as systemic lupus
 CC erythematosus and arthritis, cell depletion, neurological disorders,
 CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
 CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
 CC aplastic anaemia, myocardial infarction and stroke. The sequence
 CC presented is the human caspase recruitment domain-5 (CARD-5) protein.
 XX
 SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 23; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAILDALENLTABELKKFKLLSVLPREGYGRIPRGALLSMDALDITDKLVSFY 60
 DB 1 MGRARDAILDALENLTABELKKFKLLSVLPREGYGRIPRGALLSMDALDITDKLVSFY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGTQAPPOSAAPKPLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGTQAPPOSAAPKPLHFDQHRA 120
 QY 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 DB 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 6
 AA017854
 ID AA017854 standard; Protein; 195 AA.
 AC AA017854;
 XX
 XX 20-AUG-2002 (first entry)
 DT
 DE Pyrin domain containing protein Pycard.
 XX
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antineumatic; antiasthmatic;
 KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX
 OS Unidentified.
 XX
 XX WO200240668-A2.
 PN
 PD 23-MAY-2002.
 XX
 XX 30-OCT-2001; 2001WO-EP12545.
 PF
 XX 15-NOV-2000; 2000DE-1056687.
 PR
 PR 30-NOV-2000; 2000DE-1059595.
 PR

Db 121 ALIARVTNVWLLDALYKGLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
Db 181 LRESQSYLVEDLERS 195
RESULT 8
ID AAG74647
XX AAG74647; standard; Protein; 205 AA.
AC AAG74647;
XX
DT 03-SEP-2001 (first entry)
XX Human colon cancer antigen protein SEQ ID NO:5411.
DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
KW Homo sapiens.
XX
XX WO200122920-A2.
PN 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26524.
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI WPI; 2001-235357/24.
XX N-PSDB; AAH34052.
DR Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PS Claim 11; Page 7035-7036; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 205 AA;
Query Match 100.0%; Score 990; DB 22; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAILDALNLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
Db 11 MGRARDAILDALNLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 70

QY 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFIDQHRA 120
Db 71 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFIDQHRA 130
QY 121 ALIARVTNVWLLDALYKGLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDILLQA 180
Db 131 ALIARVTNVWLLDALYKGLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDILLQA 190
QY 181 LRESQSYLVEDLERS 195
Db 191 LRESQSYLVEDLERS 205
RESULT 9
ID AAE00594 standard; Protein; 176 AA.
XX
AC AAE00594;
XX
DT 02-JUL-2001 (first entry)
XX
DE Alternatively spliced form of human TMS1 protein (lacking exon2).
XX Human; target of methylation-induced silencing-1; TMS1; Cytostatic;
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast.
XX
OS Homo sapiens.
XX WO200129235-A2.
PN 26-APR-2001.
XX
PD 18-OCT-2000; 2000WO-US28747.
XX
PF 18-OCT-1999; 99US-0159975.
XX (UYEM-) UNIV EMORY.
PA Vertino PM;
PI
XX WPI; 2001-290922/30.
DR N-PSDB; AAD03906.
XX Novel gene TMS1, transcriptionally silenced due to increased
PT methylation useful for identifying subject at risk of developing tumor
PT characterized by abnormal methylation, for treating cancer by inducing
PT apoptosis -
XX
PS Claim 85; Page 123; 124pp; English.
XX The invention relates to identification of target of methylation-induced
CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC as tools for diagnosing and treating a subject at risk of developing
CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is alternatively spliced form of human target of
CC methylation-induced silencing-1 (TMS1) protein lacking exon2.
XX
SQ Sequence 176 AA;
Query Match 88.2%; Score 873.5; DB 22; Length 176;
Best Local Similarity 90.3%; Pred. No. 2.3e-88;
Matches 176; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60
 DB 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDQHRA 101
 QY 121 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPMSKMRKLFSTPAWNWTKDILLQA 180
 DB 102 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPMSKMRKLFSTPAWNWTKDILLQA 161
 QY 181 LRESQSYLVEDLERS 195
 DB 162 LRESQSYLVEDLERS 176

RESULT 10
 AAB43675
 ID AAB43675 standard; Protein; 190 AA.
 AC AAB43675;
 DT 08-FEB-2001 (first entry)
 XX Human cancer associated protein sequence SEQ ID NO:1120.
 DE Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaschmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX Homo sapiens.
 XX WO200005350-A1.
 XX 21-SEP-2000.
 PF 08-MAR-2000; 2000WO-US05882.
 XX 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Ruben SM;
 PI WPI; 2000-587533/55.
 DR N-PSDB; AAC77894.
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -
 PS Claim 11; Page 1731-1732; 2352pp; English.
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaschmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of

CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX Sequence 190 AA;
 SQ

Query Match 76.7%; Score 759; DB 21; Length 190;
 Best Local Similarity 95.6%; Pred. No. 1.2e-75;
 Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60
 DB 11 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 70
 QY 61 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDQHRA 120
 DB 71 LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDQHRA 130
 QY 121 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPMSKMR 160
 DB 131 ALIARVTNVVWLLDALYKGLVLTDEQYQAVR--PSPPTQAR 168

RESULT 11
 AAE00592
 ID AAE00592 standard; Protein; 193 AA.
 XX AAE00592;
 AC AAE00592;
 DT 02-JUL-2001 (first entry)
 XX Mouse target of methylation-induced silencing-1 (TMS1) protein.
 DE Mouse; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; Cpg island;
 KW caspase-recruiting domain; CARD; cancer; breast.
 XX Mus musculus.
 XX WO200129235-A2.
 XX 26-APR-2001.
 PD 18-OCT-2000; 2000WO-US28747.
 PF 18-OCT-1999; 99US-0159975.
 XX (UYEM-) UNIV EMORY.
 PA Vertino PM;
 PI WPI; 2001-290922/30.
 DR N-PSDB; AAD03904.
 XX Novel gene TMS1, transcriptionally silenced due to increased
 PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis -
 XX Claim 85; Page 120; 124pp; English.
 XX The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or

CC abnormally low levels of TMS1 expression products. Unique fragments of
 CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by
 CC increasing TMS1 polypeptide level to an above normal level. The CpG
 CC island region of TMS1 or its fragments are used to study the methylation
 CC patterns apart from any coding region contained in it.
 CC The present sequence is mouse target of methylation-induced silencing-1
 CC (TMS1) protein.
 XX Sequence 193 AA;
 SQ

Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGHLFDQHRA 120
 DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKESGAVAAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNVEMLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNWTKDQLLQA 180
 DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMKLFSFVPSNWLTKDQLLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193

RESULT 12
 AAB20086
 ID AAB20086 standard; Protein; 193 AA.
 AC AAB20086;
 DT 23-APR-2001 (first entry)
 DE Mouse CARD-5 protein.
 KW CARD-5; caspase recruitment domain; mouse; cancer; infection;
 KW autoimmune disease; neurological disease; haematological disease;
 KW immune disease; inflammation; antitumour; antiseptic;
 KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
 KW gene therapy.
 XX Mus sp.
 XX Key Location/Qualifiers
 FT Domain 110..179
 FT /note= "CARD"
 XX WO200100826-A2.
 XX 04-JAN-2001.
 XX 28-JUN-2000; 2000WO-US17691.
 XX 28-JUN-1999; 99US-0340620.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Bertin J;
 XX WPI; 2001-061973/07.
 XX N-PSDB; AAF30008.
 XX Isolated intracellular proteins predicted to be involved in regulating
 XX caspase activation are used for diagnosis and treatment of e.g. cancer,
 XX viral infections, autoimmune diseases, neurological diseases and
 XX haematological disorders -

XX Claim 9; Fig 19; 208pp; English.
 PS The present sequence is that of mouse caspase recruitment domain 5
 XX (CARD-5), an intracellular protein predicted to be involved in
 CC regulating caspase activation. The sequence is predicted from an
 CC isolated cDNA clone (see AAF30008). It shows 71.8% amino acid
 CC identity to human CARD-5 (see AAB20085). Methods of diagnosing
 CC and treating patients suffering from a disorder associated with an
 CC abnormal level or rate of apoptotic cell death, abnormal activity
 CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
 CC necrosis factor receptor complex or abnormal activity of a caspase
 CC involve administering a compound that modulates the expression or
 CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
 CC therapy methods. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD-3,
 CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
 CC cell survival and cell growth. They can also be used to screen
 CC drugs or compounds that modulate their activity or expression and
 CC to treat disorders associated with insufficient or excessive
 CC production of CARD-3, -4, -5 or -6 protein, or production of an
 CC aberrant protein.
 XX Sequence 193 AA;
 SQ

Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGHLFDQHRA 120
 DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKESGAVAAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNVEMLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNWTKDQLLQA 180
 DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMKLFSFVPSNWLTKDQLLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193

RESULT 13
 AAU99352
 ID AAU99352 standard; Protein; 193 AA.
 AC AAU99352;
 XX 07-OCT-2002 (first entry)
 DE Mouse caspase recruitment domain-5 (CARD-5) protein.
 XX Mouse; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX Mus sp.
 XX Key Location/Qualifiers
 FH

Domain 110..193
/label= CARD_domain
WO200244354-A2.
06-JUN-2002.
29-NOV-2001; 2001WO-US44894.
01-DEC-2000; 2000US-0728721.
24-APR-2001; 2001US-0841879.
(MILL-) MILLENIUM PHARM INC.
Bertin J;
WPI: 2002-557538/59.
N-PSDB; ABX87964.
Novel isolated murine or human caspase recruitment domain (CARD)-5 polypeptide, useful for treating immune disorders such as Hashimoto's thyroiditis, graft rejection, allergy, glomerular nephritis, tuberculosis -
Claim 22; Fig 1; 100pp; English.
The invention discloses the isolated polypeptides, and encoding nucleic acids, of murine and human caspase recruitment domain (CARD)-5. Caspases (cysteine aspartate-specific proteinases) are central to the apoptotic program and responsible for the degradation of cellular proteins that lead to the morphological changes seen in cells undergoing apoptosis. Caspases interact with other caspases via their CARDS and different subtypes of CARDS may confer binding specificity. CARD-5 is an intracellular protein that is predicted to be involved in regulating caspase activation. CARD-5 activates the nuclear factor-kappa B (NF-kappaB) transcription factor pathway and binds the CARDS of caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5 activity and NF-kappaB activation, regulate cell growth and cell death and be used in gene therapy. The CARD-5 polypeptides are useful for identifying compounds which bind to them and modulate their activity and for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides, nucleic acids, antibodies and modulators of CARD-5 expression or activity can be used to treat immune disorders such as chronic inflammatory diseases and disorders, Hashimoto's thyroiditis, graft rejection, sarcoidosis, atopic conditions (such as asthma and allergy), glomerular nephritis, human immunodeficiency virus (HIV) and bacterial infections (including tuberculosis and lepromatous leprosy) and in screening and detection assays. Modulators of CARD-5 activity or expression are also useful for treating autoimmune disorders, such as systemic lupus erythematosus and arthritis, cell depletion, neurological disorders, such as Alzheimer's disease, Parkinson's disease and spinal muscular atrophy, haematologic diseases, such as myelodysplastic syndrome and aplastic anaemia, myocardial infarction and stroke. The sequence presented is the mouse caspase recruitment domain-5 (CARD-5) protein.
Sequence 193 AA;
Query Match 69.1%; Score 684; DB 23; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-67;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
QY 1 MGRARDAILDALENLTAEELKFKKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRARDAILDALENLSGDELKFKKLLTVQLREGYGRIPRGALLQMDALDITDKLVSY 60
QY 61 LETYGAELTANVLNDRMGLQEMAGLOQAATHQSGAAPAGIQAPPOSAKPGLFHIDQHRA 120
DB 61 LESYGLTMTVLNDRMGLQELAEQLQ-TTKESGAVAAASVPAQSTARTG-HFVDQHQ 118
QY 121 ALIARVTNVEILLDALYKGVLTDEQYQAVRAPPTNPSKVKLFSTFPANWTCKLLLOA 180
DB 119 ALIARVTEVDGLDALHGSVLTEGQYQAVRAETTSQDKMKLFSFVPSWNLCKDSLLQA 178

QY 181 LRSQSYLVLEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193
RESULT 14
AAE00593
ID AAE00593 standard; Protein; 171 AA.
XX
AC AAE00593;
XX
DT 02-JUL-2001 (first entry)
XX
DE Rat target of methylation-induced silencing-1 (TMS1) partial protein.
XX
KW Rat; target of methylation-induced silencing-1; TMS1; cytostatic;
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT Misc-difference 36
FT /note= "Encoded by AGY"
XX
PN WO200129235-A2.
XX
PD 26-APR-2001.
XX
PF 18-OCT-2000; 2000WO-US28747.
XX
PR 18-OCT-1999; 99US-0159975.
XX
PA (UDEM-) UNIV EMORY.
XX
XX Vertino PM;
XX
XX WPI; 2001-290922/30.
DR N-PSDB; AAD03905.
XX
XX Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor characterized by abnormal methylation, for treating cancer by inducing apoptosis -
XX
PS Claim 85; Page 121; 124pp; English.
XX
CC The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it.
CC The present sequence is rat target of methylation-induced silencing-1 (TMS1) partial protein.
XX
SQ Sequence 171 AA;
Query Match 58.2%; Score 576; DB 22; Length 171;
Best Local Similarity 67.1%; Pred. No. 2e-55;
Matches 116; Conservative 18; Mismatches 37; Indels 2; Gaps 2;
QY 23 FKLLSVPLREGYGRIPRGALLSMDALDITDKLVSVFYLETYGAELTANVLNDRMGLQEMA 82
DB 1 FKILLTAPVREGYGRIPRGALLQMDPDLITDKLVXYLEGYGLTMTVLNDRMGLQELA 60

QY 83 GQLQAATHOGSGAAPAGIOAPPQSAAKPGLHFIHQHRAALIARVTNVWLLDALYGVLT 142
Db 61 EQ-LKIMEB-SGAVATATSPVAGTARTE-HFVDQHRQALIARVTEVDGLLDALYGVLT 118
QY 143 DEQYQAVRAEPTNPSPKRLFSFTTPAWNWTCKDLLQALRESQSYLVDELRS 195
Db 119 EGOYQAVRAETTNQNRKRLFSAPAWNLTCKNLFLEALRQTPYLVTLEQS 171

RESULT 15
AAG74648
ID AAG74648 standard; Protein; 158 AA.
XX AAG74648;
XX 03-SEP-2001 (first entry)
DT Human colon cancer antigen protein SEQ ID NO:5412.
DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX Homo sapiens.
XX WO200122920-A2.
PN 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26524.
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI WPI; 2001-233357/24.
DR N-PSDB; AAH34053.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 11; Page 7036-7037; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX SQ Sequence 158 AA;

Query Match 57.0%; Score 564.5; DB 22; Length 158;
Best Local Similarity 75.8%; Pred. No. 3 4e-54;
Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;
QY 4 ARDAILDALENLTABELKKFLLKLVSVFYLET 63
|||||

Db 13 ARDAILDALENLTABELKKF-----KLVSVFYLET 41
QY 64 YGAELTANVLRDMGLQEWAGOLQAATHOGSGAAGPAGIOAPPQSAAKPGLHFIHQHRAALI 123
Db 42 YGAELTANVLRDMGLQEWAGOLQAATHOGSGAAGPAGIOAPPQSAAKPGLHFIHQHRAALI 101
QY 124 ARVTNVEWLLDALYGVLTDEQYQAVRAEPTNPSPKMR 160
Db 102 ARVTNVEWLLDALYGVLTDEQYQAVR--PSPPTQAR 136

Search completed: January 29, 2004, 13:45:54
Job time : 9.48584 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 4.80296 Seconds
(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDALDALENLTAEEL.....LLQLALRBSQSYLEVDLERS 195
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	25.3	1192	T17255	hypothetical prote
2	87	8.8	323	T49723	hypothetical prote
3	86.5	8.7	372	S23326	gene ML2.2 protein
4	86.5	8.7	439	S75545	hypothetical prote
5	85	8.6	237	E55210	heti protein - Ana
6	85	8.6	1488	AH0615	cell division prot
7	84	8.5	292	C64699	proteinase IV (EC
8	83.5	8.4	542	A84229	sensory histidine
9	83	8.4	455	JC6051	trypsin-like prote
10	83	8.4	545	A87136	hypothetical prote
11	83	8.4	877	JN0772	glucan endo-1,3-be
12	82.5	8.3	446	S16308	translation elonga
13	82.5	8.3	522	D82195	phage replication
14	82	8.3	222	F71819	probable proteinas
15	82	8.3	788	A72330	phenylalanine-trna
16	81	8.2	792	AB0704	pyruvate, water di
17	80.5	8.1	718	B81058	guanosine-3',5'-bi
18	80.5	8.1	725	E81819	guanosine-3',5'-bi
19	80.5	8.1	2117	T36180	CDA peptide synthe
20	80	8.1	556	F87297	feruloyl-CoA synth
21	80	8.1	1233	T40059	chromosome segrega
22	79	8.0	350	E87714	hypothetical prote
23	79	8.0	455	C91142	serine endoprotein
24	79	8.0	455	F85987	serine endoprotein
25	79	8.0	860	G82110	alanyl-tRNA synthe
26	79	8.0	1486	C64832	cell division prot
27	79	8.0	1486	G90754	kinesin-like cell
28	79	8.0	1486	E85618	hypothetical prote
29	79	8.0	1534	JH0228	cell division prot

30	78.5	7.9	322	2	F87549	type IV secretion
31	78.5	7.9	346	2	A48470	translation elonga
32	78.5	7.9	424	2	AC2975	oxidoreductase ord
33	78.5	7.9	424	2	H98307	probable oxidoredu
34	78.5	7.9	699	2	T18984	hypothetical prote
35	78.5	7.9	956	2	B71468	probable insulinas
36	78.5	7.9	968	2	T51933	kinesin motor prot
37	78.5	7.9	4684	2	A59404	plectin [imported]
38	78	7.9	325	2	A97482	hypothetical prote
39	78	7.9	325	2	A12699	conserved hypothet
40	78	7.9	335	2	C75619	histidinol-phospha
41	78	7.9	351	2	A83419	phage host specifi
42	78	7.9	366	2	G84849	probable actin [im
43	77.5	7.8	228	2	C87467	ABC transporter, A
44	77.5	7.8	542	2	JN0661	heat shock protein
45	77.5	7.8	542	2	S32106	groEL protein - La

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T17255
R/Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A/Reference number: Z18722
A/Accession: T17255
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1192 <KOE>
A/Cross-references: EMBL:AL117470
A/Experimental source: adult uterus; clone DKFZp586O1822
C/Genetics:
A/Note: DKFZp586O1822.1

Query Match	25.3%	Score	250;	DB 2;	Length	1192;
Best Local Similarity	44.5%	Pred. No.	1.7e-13;			
Matches	65;	Conservative	23;	Mismatches	48;	Indels 10; Gaps 4;
QY	54	DKLVS-FYLETYGAEILTANV--LRDMGLQEMA-----GOLQAATHQGSAAPIQAPPOS	106			
Db	1036	DQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKFGDLMPAT---TLIPPARIAVPSPL	1092			
QY	107	AAKPGHFIIDQHRALIAIARTVWMLLDALYKVLTDQYQAVRAPPTWPSQMKLFSFT	166			
Db	1093	DAPQLLHFVDQYREQLIARTVTSVEVVDLKHGQVLSEQYERVLAEINTRPSQMRKLFSL	1152			
QY	167	PAWNWTCKDLLQALRESQSYLEVDL	192			
Db	1153	QSWDRCKDGLYQALKETHPHLMEL	1178			

RESULT 2

T49723
hypothetical protein B23L21.360 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49723
R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49723
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-323 <SCH>
A/Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.360
A/Experimental source: BAC clone B23L21; strain OR74A
C/Genetics:
A/Gene: NCSP:B23L21.360

A;Map position: 6
A:Introns: 183/1: 229/2

Query Match	8.8%;	Score 87;	DB 2;	Length 323;
Best Local Similarity	20.7%;	Pred. No. 5.1;		
Matches 44;	Conservative 41;	Mismatches 80;	Indels 48;	Gaps 9;
Qy	6	DAILDALENLTAEELKKFKLLKLVPLREGYGRIPRGALLSMDALDLTDLKLVSYFLSTYG	65	
Db	119	DNHFDDLKVQGEDFEWFDEQVLSLP-----SSPTVIERFATENMAKHVIISRG	168	
Qy	66	AELTANVLRDMGLQEMAGQIOATHQSGGAAPGAGIQAPPOSAAKPGUHLFDHQRAALIAR	125	
Db	169	-----EMSVR-MFGVVVA--HAGSGTILEVLRLQVLVVPNPETLMDNHOAEL----	213	
Qy	126	VTNVEWLDDALYGVKV--LTDEQYQAVRAEPTNPCKMKLFSFTB-----AANNW	171	
Db	214	AQELSTGDDAVYGLGKLKLTAIARSLVLAQGPLKLDLPYPSPPPFVDPDSERVTLFDW	273	
Qy	172	---TC-----KOLLQALRESQVYLVEDLERS	195	
Db	274	MTLTCYPDELRRKOOHLODLNRNVEKSFQORELEQA	306	

```

RESULT 3
S23326
gene M16.2 protein precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S23326
R:Bessen, D.E.; Fischetti, V.A.
Infect. Immun. 60, 124-135, 1992
A:Title: Nucleotide sequences of two adjacent M or M-like protein genes of
A:Reference number: S23325; MUID:92104662; PMID:1370269
A:Accession: S23326
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <BES>
A:Cross-references: EMBL:X61276; NID:g47369; PIDN:CAA43582.1; PID:g47371
C:Superfamily: M5 protein

```

Query Match	8.7%; Score 86.5; DB 2; Length 372;
Best Local Similarity	26.8%; Pred. No. 6.7;
Matches 53; Conservative	30; Mismatches 78; Indels 37; Gaps 9;

QY	12	L E N U T A B E L K F K L L S V P R E C Y G R I P R G A L L S M D A L - - - - D L T D K L V S F Y L E T Y G A	66
DB	78	L E K I N A E E E K K L E A I N K E L N E N Y K L Q D G - - - - I D A L E K E K E D L T T L A K T T K E N E I S	133
QY	67	E L T A N V L - R D W G L O E M A G O L Q A T H O G S G A A P A G I O A P P Q - - S A A K P G L - H F I D O H R A A L	122
DB	134	E A S R K G L S R D L E A S R T A K K E L A E A K H Q K L E A N K K L T E G N Q V S E A S R K G L S N D L E A S R A A -	192
QY	123	I A R V T N V E W L L D A L Y G K V L T D E Q - - - - Y Q A V R A E P T N P S K M R K L F S P T A W N W T C K L L	177
DB	193	- - - - - K K E L A Y Q K L E T D H Q L E A K H O K L E A D Y Q V S E T S R K G L S - - - - - R D - - -	234
QY	178	L Q A L R E S Q S Y L V E D L E R S	195
DB	235	L E A S R E A N K K V T S E L T O A	252

RESULT 4
S75545
hypothetical protein alr0806 - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S75545
R;Kotani, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; H
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamad
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium

S.
A; Reference number: S74322; MUID: 97061201; PMID: 8905231

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-439 <RAN>
A/Cross-references: EMBL:D90911; GB:A5001339; NID:G1653083; PIDN:BAA18106.1; PID:G1653190
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: glycolate oxidase chain gicD

	Query Match	8.7%;	Score 86.5;	DB 2;	Length 439;
	Best Local Similarity	24.0%;	Pred. No. 8.2;		
	Matches	42;	Conservative	22;	Mismatches 60; Indels 51; Gaps 8;
QY	31	PLREGYGR---	TPRGALLSMDALD	LDLTKLVSFYLETY	GAEALTANVLMDGLQEWAGLQA 87
Db	194	PVHNGVTGTG	ITETITLPTPAL	PWEAIVSF-----	TNLSAIAFAQN 237
QY	88	ATHQ--GSGA	APAGIQAPP-----	QSAAKPGLHFIDQ	HRAALIAVRTNVW---LL 133
Db	238	LAHQDIVSK	ETISIQADPIQ	POYFSSLKSYQ	PGAHV-----MVISLDMIAFTQLA 290
QY	134	DALYKVLTD	EYQAVRAEPTNP	SKMRKLFSP	TPANNWTKDLLLQALRESQYL 188
Db	291	KASKGIIPEO	-----DPSG	KKINLIEF-----	NNHNTTLARAVDPSLYL 334

RESULT 5

E55210
 A:het1 protein - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 15-Oct-1999
 C:Accession: E55210
 J;Black, T.A.; Wolk, C.P.
 R. Bacteriol. 176, 2282-2292, 1994
 A:Title: Analysis of a Het - mutation in Anabaena sp. strain PCC 7120 implicates a second
 A:Reference number: A55210; MUID:94209228; PMID:8157596
 A:Accession: E55210
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-237 <BLA>
 A:Cross-references: GB:L22883; NID:G441119; PIDN:AAA22003.1; PID:G441121
 C:Genetics:
 A:Gene: het1
 A:Start codon: TTG

Query Match	8.6%;	Score 85;	DB 2;	Length 237;	
Best Local Similarity	23.7%;	Pred. No. 5.2;			
Matches	49;	Conservative 32;	Mismatches 72;	Indels 54;	Gaps 12;
QY	14	NLTAECLKFKLLSVLPREGVGRTPRGALLSMDALDUTDKLV---SFYLETYGAELTA	70		
DB	12	NLT---LLSDEVHLWRIPLDQ-----PESQLQDLAATLSDELARANRFVFPERRRFTA	63		
QY	71	--NVLDMGLQEWAGOLOAATHQGSGAAPAGIQAPQSAKP-----GLHFIDQHR	119		
DB	64	GRGILRSI----LGGYL-----GVPEQCVKFDYESRGKPTILGRFAESGLLFLNLSHS	111		
QY	120	AALIARVTN-----VFWL-----LDALYKVLVTDEQYQVAVRAEPTNPSKMKLKFST	166		
DB	112	QNALCAVNYTRQIGIDLYLRPTSDLESIAKRFFLPREYELLSLP-DEQKQIKIFYR-	169		
QY	167	PAWNWTCKDLLLQAALRESQSLYLEDLE	193		
DB	170	----WTCKEAYLKATGDGIAGL-EETTE	191		

RESULT 6

cell division protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (str. ...)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence revision 09-Nov-2001 #text change 18-Nov-2002

C:Accession: AH0615
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0615
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1488 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05394.1; PID:g16502156; GSPDB:GN00176
C:Genetics:
A:Gene: mukB
C:Superfamily: Escherichia coli mukB protein

Query Match 8.6%; Score 85; DB 2; Length 1488;
Best Local Similarity 24.0%; Pred. No. 50;
Matches 42; Conservative 33; Mismatches 56; Indels 44; Gaps 11;
QY 3 RARDALD-----ALENLTAEELKKPKLLSVP-----LREGYGRIPRGALLS- 46
Db 638 RREALVERDEVGARKNAVDEIER-----LSOPGGAEDQRNALAERFG-----GVLLSE 688
QY 47 -MDALDLTKLVSYLETYCAELTANVLDRMDGLQEMAGLOAATHOGSGAAPAG---IOA 102
Db 689 IYDDVSLED--APYPSALYGPSRHAIVPD--LSQIAEQLEGITD-----CPEDLYLIEG 739
QY 103 PQSAAKPGIHFDHRAALIAKRVNVEWLLD-----ALYKVLTDQVQAVRAE 152
Db 740 DRQS-FDDSVFSDLEKAVVVKIADROWRYRFPSPSLPIFGRAARENRIESLHAE 793

RESULT 7
C64699
proteinase IV (EC 3.4.-.-) - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 15-Aug-1997 #text_change 29-Sep-1999
C:Accession: C64699
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: C64699
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-292 <TOM>
A:Cross-references: GB:AE000644; GB:AE000511; NID:g2314609; PIDN:AAD08477.1; PID:g231461
C:Superfamily: proteinase IV MTH806
C:Keywords: hydrolase

Query Match 8.5%; Score 84; DB 2; Length 292;
Best Local Similarity 23.6%; Pred. No. 8.2;
Matches 41; Conservative 30; Mismatches 61; Indels 42; Gaps 7;
QY 42 GALLSMDALDL-TDKL-----VSFYLETYCAELTANY-----LRDMLQEMAGLOA 87
Db 57 GAIFSTEDFDKEVDKILKTPSIKGVLLIDSPGCAVSASVELSEKIADL-KQMPVLAVA 115
QY 88 ATHQSGAAPAGTQAPQSAKPGIHFIDHRAALIAKRVNVEWLLD-----TNVEWLLDALYKVL 141
Db 116 RGVMSGSGYVYAGQASVEVASK-----ASLIGSIGVIFSGANVENLNKV----- 160
QY 142 TDQYQAVRAEPTNPSPKRLFSFTPAWNWTKDILLQALRESQSYLVEDLERS 195
Db 161 -----GVATQGVHAGEYKEIGFTTRAKPNKEDFLQNLVNEQYQMFVNDVAKA 208

RESULT 8
Query Match 8.4%; Score 83; DB 2; Length 455;
Best Local Similarity 21.9%; Pred. No. 17;
Matches 55; Conservative 34; Mismatches 76; Indels 86; Gaps 10;

A84229
sensory histidine protein kinase homolog [imported] - *Halobacterium* sp. NRC-1
C:Species: *Halobacterium* sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84229
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leichauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: A84229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-542 <STO>
A:Cross-references: GB:AE004437; NID:g10580295; PIDN:AAG19197.1; GSPDB:GN00138
C:Genetics:
A:Gene: aisoQ

Query Match 8.4%; Score 83.5; DB 2; Length 542;
Best Local Similarity 28.2%; Pred. No. 19;
Matches 37; Conservative 18; Mismatches 53; Indels 23; Gaps 6;
QY 37 GRIPGALLSMDAL--DLTDKLVSFYLE-----TYGAELTANVLRDM--GLQEMAGLOA 88
Db 40 GRIEYANEAAARDVIGADARDLVGTFFQAPWFTHDQVQAAVREQVAAAVGEASQF-TA 98
QY 89 THQSGAAPAGTQAPQSAKPGIHFIDH-----AAALIAKRVNVEWLLD--- 134
Db 99 THRTSGGTATVELELQMPAPAVDGDHDAFVAVVVVGRRAAESTAADVESALDAVQ 158
QY 135 ALYKVLTDQ 145
Db 159 ALYATDIDQ 169

RESULT 9
JC6051
trypsin-like proteinase (EC 3.4.21.-) hhoA precursor - *Escherichia coli* (strain K-12)
C:Species: *Escherichia coli*
C:Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 01-Mar-2002
C:Accession: JC6051; D65115
R:Bass, S.; Gu, Q.; Christen, A.
J. Bacteriol. 178, 1154-1161, 1996
A:Title: Multicopy suppressors of Prc mutant *Escherichia coli* include two HtrA (DegP) pr A:Reference number: JC6051; MUID:96165273; PMID:8576052
A:Accession: JC6051
A:Molecule type: DNA
A:Residues: 1-455 <BAS>
A:Cross-references: GB:U15661; NID:g558911; PIDN:AAC43992.1; PID:g558912
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D65115
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-455 <BLAT>
A:Cross-references: GB:AE000402; GB:U00096; NID:g1789619; PIDN:AAC76266.1; PID:g1789629; A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: hhoA
C:Superfamily: Helicobacter serine proteinase
C:Keywords: hydrolase; serine proteinase
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-455/Product: trypsin-like serine protease hhoA #status predicted <MAT>
F;109,139,214/Active site: His, Asp, Ser #status predicted

Query Match 8.4%; Score 83; DB 2; Length 455;
Best Local Similarity 21.9%; Pred. No. 17;
Matches 55; Conservative 34; Mismatches 76; Indels 86; Gaps 10;

QY 11 ALNLTAEELKFKLLKLVPLREGYGRIPRG-ALLSMDALDITDKLVSF-----YLET 63
 Db 217 ALLNLNGE-----LGINTAILAPGGSGVIGFAPSPNNWARTLAQOLIDFGSIKRLGCI 271
 QY 64 YGAELTANVLRDMGLQEMAGQIQAATHQSGGAAPAGIQIA-----102
 Db 272 KGTMSADIKAFLNLDVQRFVSEVLPGSGSAKAGVKGAGDIITSLNGKPLNSFAELRSR 331
 QY 103 -----PPQSAKPGCLHFIHQHRAALIAVTV-----NVEMLDALYKVLTDQY---146
 Db 332 IATTEPGTKVILGGLL---BNGKPLEVEVLTDTSTSSASAEMITPALEGATILSDGQLKDG 388
 QY 147 -----QAVRAEP-----TNPSKWRKLFSTPAWNTCKDILLQ 179
 Db 389 GGIKIDIVKSGSPAAQAGLQKDDVIIGNVRDRVNSIAEMRKVLAAPKPA-----IIALQ 442
 QY 180 ALRESQS-YLV 189
 Db 443 IVRGNESIYLL 453
 RESULT 10
 A: Molecule type: protein
 A: Residues: 1-189
 A: Cross-references: GB:AL450380; NID:gl3093526; PIDN:CAC30769.1; GSPDB:GN00147
 C: Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C: Accession: A87136
 R: Cole, S.T.; Eglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A: Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
 A: Title: Massive gene decay in the leprosy bacillus.
 A: Reference number: A86909; PMID:21128732; PMID:11234002
 A: Accession: A87136
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-189
 A: Cross-references: GB:AL450380; NID:gl3093526; PIDN:CAC30769.1; GSPDB:GN00147
 C: Genetics: A87136
 C: Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 Query Match 8.4%; Score 83; DB 2; Length 545;
 Best Local Similarity 24.4%; Pred. No. 22;
 Matches 41; Conservative 15; Mismatches 50; Indels 62; Gaps 7;
 QY 51 DLTDLKLVSFY-----LETYGAELTA-NVLRDMGLQ-----EM 81
 Db 113 DARDRAIRYQLEERFLALGGYSAEAGRICASGLPERVLVQQLCTLSGGQRRVEL 172
 QY 82 AQQLQATHQSGGAAPAG-----IQAPPOSAKPGCLHFIHQHRAALIA 124
 Db 173 ARILFAASRAGTCASGGTLLDDEPTNHLDAADSLGWLDRFLRSHTGGLVWISHNVLLIA 232
 QY 125 RVTNVEMLDALYKVL-----LTDEQYQAVRAEPTNPSK 158
 Db 233 AVNVRVWFLDAVLGKVDVYVMGMWYKYLSRATDEQRR--RRERVNAER 278
 RESULT 11
 JN0772
 Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) precursor - Bacillus circulans (strain
 N: Alternate names: beta-1,3-glucanase
 C: Species: Bacillus circulans
 C: Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
 C: Accession: JN0772; P0613
 R: Yamamoto, M.; Aono, R.; Horikoshi, K.
 Biosci. Biotechnol. Biochem. 57, 1518-1525, 1993
 A: Title: Structure of the 87-kDa beta-1,3-glucanase gene of Bacillus circulans IAM1165 a
 A: Reference number: JN0772; MUID:94033858; PMID:7764221
 A: Accession: JN0772

A: Molecule type: DNA
 A: Residues: 1-877 <YAM>
 A: Cross-references: GB:D17519; NID:g601877; PIDN:BA04469.1; PID:g601878
 A: Accession: P0613
 A: Molecule type: protein
 A: Residues: 39-58 <YAS>
 C: Genetics:
 A: Gene: bglH
 C: Keywords: glycosidase; hydrolase; polysaccharide degradation
 F: 1-38/Domain: signal sequence #status predicted <SIG>
 F: 39-877/Product: glucan endo-1,3-beta-D-glucosidase #status predicted <MAT>
 Query Match 8.4%; Score 83; DB 2; Length 877;
 Best Local Similarity 22.3%; Pred. No. 39;
 Matches 40; Conservative 26; Mismatches 73; Indels 40; Gaps 8;
 QY 10 DALENLTAEELKFKLLKLVPLR-----EGYGRIPRGALLSMDALDITDKLVSYLE-- 62
 Db 471 DRAQNVFVQD-GKLNKALNEPKSPQDPSPRYAQYSSGKINTKDHFSLYKGRVDFAKLIP 529
 QY 63 -----TYGAELTANVLRDMGLQEMAGQIQAATHQSGGAAPAGIQAPQS 106
 Db 530 TONGIWPALWMLPQDNVYGTWASSG---EIDWMAKGRDPST---SGAVHGGQWPTNR 583
 QY 107 AAKPGLHF-----IDQHRAALIAVTVNVEMLDA-LYGVKVLTDQYQAVRAEPTNPS 157
 Db 584 YLSGEYHFPPEGOTFANDYHYVYVWVEEDNIKWYVDGKFFKVTROQWYSA--AAPNNPN 640
 RESULT 12
 S16308
 Translation elongation factor eEF-1 alpha chain - Stylonychia lemnae
 C: Species: Stylonychia lemnae
 C: Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
 C: Accession: S16308
 R: Bierzbaum, P.; Doenhoff, T.; Klein, A.
 Mol. Microbiol. 5, 1567-1575, 1991
 A: Title: Macromolecular and micronuclear configurations of a gene encoding the protein synt
 A: Reference number: S16308; MUID:92157884; PMID:1840642
 A: Accession: S16308
 A: Molecule type: DNA
 A: Residues: 1-446 <BIE>
 A: Cross-references: EMBL:X57926; NID:gi10139; PIDN:CAA41001.1; PID:gl10140
 A: Note: the authors translated the codon GAT for residue 307 as Gly
 C: Genetics:
 A: Gene: efaA
 A: Superfamily: SGC5
 C: Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
 C: Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis
 F: 8-156/Domain: translation elongation factor Tu homology <ETU>
 F: 14-21/Region: nucleotide-binding motif A (P-loop)
 F: 153-156/Region: GTP-binding NKXD motif
 Query Match 8.3%; Score 82.5; DB 2; Length 446;
 Best Local Similarity 21.4%; Pred. No. 19;
 Matches 44; Conservative 40; Mismatches 81; Indels 41; Gaps 8;
 QY 8 ILDALENLTAEELKFKLLKLVPLREGY-----GRIPRGAL-----LSMDALDL 52
 Db 216 LIDALDAL--DQKRPDKPLRLPLQDYKIGITVPVGRVETGLLPGWVLTFAPMNI 273
 QY 53 TDKLVSYLETYGAELTA-----NV---LRDMGLQEMAGQIQAATHQSGGA-----AP 97
 Db 274 TTECKS--VEMHESLTAEPDGNVGFVKNLSVXDLRGVYASDKNDPAKDTTNFLAQ 331
 QY 98 AGIQAPPOSAKPGCLHFIHQHRAALIAVTVNVEMLDALYKVLTD-----YQAVR 150
 Db 332 VIVLHPGQIQGYAPVLDCHTAHTACKFDEIESKVDRRSGKVLSEEPKFKSGBAALVR 391
 QY 151 AEPTNPSKWRKLFSTPAWNTCKDL 176
 Db 392 MYFQKPMCVFAFNQYPPILGRFAVRDM 417

RESULT 13

D82195
 phage replication protein Cri VC1469 VC1475 [imported] - Vibrio cholerae (strain N16961)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: D82195; A82196
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.L.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: D82195
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <HE1>
 A:Cross-references: GB:AE004225; GB:AE003852; NID:99555968; PIDN:AAF94626.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 A:Accession: A82196
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <HE2>
 A:Cross-references: GB:AE004225; GB:AE003852; NID:99555968; PIDN:AAF94631.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics: <GEN1>
 A:Gene: VC1469
 A:Map position: 1
 C:Genetics: <GEN2>
 A:Gene: VC1475
 A:Map position: 1

Query Match 8.3%; Score 82.5; DB 2; Length 522;
 Best Local Similarity 24.4%; Pred. No. 23;
 Matches 38; Conservative 20; Mismatches 53; Indels 45; Gaps 7;
 QY 10 DALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFYLETYGAE 68
 DB 300 DLLESKGELELYNDSKIKQLLRDAYSTWTPKGNISYAKA-----DRLFRFYM----- 348
 QY 69 TANVLDMGLQEWAGLOQAT-HQGS-----GAAPAGIQAPQSAKPGHFDHRAA 121
 DB 349 ---TLCDRGYQELKAHSSKATLRNMRDLMAIGFSKADLQNLSEGRMP----- 394
 QY 122 LIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPS 157
 DB 395 -LAQVLNFN-----DNQRPNVYVEPVST 418

RESULT 14

F71819
 probable proteinase IV - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
 C:Accession: F71819
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: F71819
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-292 <ARN>
 A:Cross-references: GB:AE001556; GB:AE001439; NID:94155938; PIDN:AAD06917.1; PID:9415595
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: sppA
 C:Superfamily: proteinase IV MTH806

Query Match 8.3%; Score 82; DB 2; Length 292;
 Best Local Similarity 23.0%; Pred. No. 12;

Matches 40; Conservative 32; Mismatches 60; Indels 42; Gaps 7;
 QY 42 GALLSMDALDL-TDKL-----VSFYLETYGAEALTANV-----LRDMGLQEWAGLOQA 87
 DB 57 GAIFSTEDFDKEVDKILKTPSIKGVLLIDSPGGASVASVELSEKIADL-KQKMPVLA 115
 QY 88 ATHQSGGAAPAGIQAPQSAKPGHFDHRAAIIARV-----TNVWLLDALYGVLT 141
 DB 116 RGVWASGSYYAGMQASEVYASK-----ASLIGSIGVIFSSANVENLKNV----- 160
 QY 142 TDEQYQAVRAEPTNPSKMKLFSTPAWNWTKDQLLQALRESQSVLVEDLERS 195
 DB 161 -----GVATQGVHAGEYKEIGTFRANKPNEKEFLQNLVNEQYQMFVDDVAKA 208

RESULT 15

A72330
 phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
 C:Accession: A72330
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, G.; Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of Thermotoga maritima.
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: A72330
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-788 <ARN>
 A:Cross-references: GB:AE001749; GB:AE000512; NID:94981346; PIDN:AAD35904.1; PID:9498135
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0822
 C:Superfamily: phenylalanine-tRNA ligase beta chain
 C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 8.3%; Score 82; DB 2; Length 788;
 Best Local Similarity 24.4%; Pred. No. 42;
 Matches 29; Conservative 19; Mismatches 57; Indels 14; Gaps 2;
 QY 6 DAILENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFYLETYG 65
 DB 87 DYVILALEGATLN-----NGLKIEPRFKGVISEGMLCSLEELGLEEKSDRVVRFDP 139
 QY 66 AEILTANVLDMGLQEWAGLOQAATHQ-----GSAAPAGIQAPQSAKPGHFDIQ 117
 DB 140 VELGVNVVEEYGLNERVLDEITPNRDCLSIIGVARELSALSGRPLNKPQDPVSFVDE 198

Search completed: January 29, 2004, 13:49:47
 Job time : 6.80296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 3.00185 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDALDALENLTAEEL.....LLQLALRESQSYLVEDLRS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	990	100.0	195	1 ASC_HUMAN	Q9ulz3 homo sapien
2	684	69.1	193	1 ASC_MOUSE	Q9epb4 mus musculus
3	250	25.3	1473	1 NAL1_HUMAN	Q9c000 homo sapien
4	213	21.5	203	1 ASC_BRARE	Q9i9n6 brachydanio
5	138	13.9	854	1 PYA5_RAT	Q63035 rattus norv
6	132	13.3	843	1 PYA5_MOUSE	Q91ws2 mus musculus
7	125.5	12.7	892	1 PYA5_HUMAN	P59044 homo sapien
8	122.5	12.4	781	1 MEVH_HUMAN	O15553 homo sapien
9	99.5	10.1	980	1 PYA3_HUMAN	Q8wx94 homo sapien
10	99	10.0	431	1 CAR8_HUMAN	Q9Y2q2 homo sapien
11	91.5	9.2	143	1 SILE_SALTY	Q9z4n3 salmonella
12	89	9.0	338	1 RUVE_THETN	Q8ran2 thermonaer
13	86.5	8.7	372	1 M22_STRPY	P50469 streptococc
14	86.5	8.7	835	1 SYFE_CORGL	Q8ngn6 corynebacte
15	86	8.7	1033	1 C1S1_MOUSE	Q8r4b8 mus musculus
16	85	8.6	237	1 HETI_ANASP	P37695 anabaena sp
17	84.5	8.5	540	1 RIK2_HUMAN	O43353 h receptor-
18	83	8.4	378	1 ACT_SCHDU	O65314 scherfelia
19	83	8.4	455	1 DEGO_ECOLI	P39099 escherichia
20	82.5	8.3	446	1 EFIA_STYLE	P25166 stylonychia
21	82	8.3	788	1 SYFB_THEMEA	Q9ws89 thermotoga
22	80	8.1	539	1 RIK2_MOUSE	P58801 mus musculus
23	80	8.1	1034	1 C1S1_HUMAN	Q96p20 homo sapien
24	80	8.1	1233	1 SMCI_SCHPO	Q94383 schizosacch
25	79	8.0	860	1 SUE_VIECH	Q56648 vibrio chol
26	79	8.0	1486	1 MYA_ECOLI	P22523 escherichia
27	78.5	7.9	346	1 EFIA_EIMBO	Q107051 eimeria bov
28	78.5	7.9	4684	1 PLEI_EIMBO	Q15149 homo sapien
29	78	7.9	366	1 ACT9_ARATH	P93738 arabidopsis
30	76.5	7.7	213	1 Y914_THEMEA	Q9x016 thermotoga
31	76.5	7.7	324	1 PARB_TREPA	O83295 treponema p
32	76.5	7.7	953	1 CAR4_HUMAN	Q9Y239 homo sapien
33	76	7.7	739	1 PURL_LISMO	Q8y6c1 listeria mo

34	75.5	7.6	444	1	TIG_AZOBR	Q9x6w7 azospirillu
35	75.5	7.6	542	1	CH60_LACLA	P37282 lactococcus
36	75.5	7.6	542	1	CH60_LACLC	Q9aep7 lactococcus
37	75.5	7.6	546	1	CH60_LACAC	Q93g07 lactobacill
38	75.5	7.6	611	1	BUR_CHICK	Q90660 gallus gall
39	75.5	7.6	1698	1	Y076_HUMAN	Q14999 homo sapien
40	75	7.6	545	1	CH60_PARDE	Q92462 paracoccus
41	75	7.6	791	1	PFSA_ECOLI	P23538 escherichia
42	74.5	7.5	401	1	ISPE_LYCES	P93841 lycopersico
43	74.5	7.5	862	1	LOXA_PHAVU	P27480 phaseolus v
44	74.5	7.5	957	1	SECA_MYCSM	P71533 mycobacteri
45	74.5	7.5	1062	1	PIA7_HUMAN	P59046 homo sapien

RESULT 1
ASC_HUMAN
ID ASC_HUMAN STANDARD; PRT; 195 AA.
AC Q9ULZ3; Q9BSZ5; Q9HBD0; Q9NXJ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apoptosis-associated speck-like protein containing a CARD (hASC)
DE (PYCARD) (Target of methylation-induced silencing 1) (Caspase
DE recruitment domain protein 5).
GN ASC OR TMS1 OR CARD5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Leukemia;
RX MEDLINE=20036508; PubMed=10567338;
RA Masumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T.,
RA Niihara N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.;
RT "ASC, a novel 22-kDa protein, aggregates during apoptosis of human
RT promyelocytic leukemia HL-60 cells";
RL J. Biol. Chem. 274:33835-33838(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fibroblast;
RX MEDLINE=20552139; PubMed=11103776;
RA Conway K.E., McConnell B.B., Bowring C.E., Donald C.D., Warren S.T.,
RA Vertino P.M.;
RT "TMS1, a novel proapoptotic caspase recruitment domain protein, is a
RT target of methylation-induced gene silencing in human breast
RT cancers";
RL Cancer Res. 60:6236-6242(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Martinon F., Hofmann K., Tschopp J.;
RT "Pycard a PYD and CARD containing molecule";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Bertin J.;
RT "CARD5 protein is a CARD/PYRIN family member that is involved in
RT apoptosis signal transduction";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Colon mucosa;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isoigai T., Sugano S.;
RT "NED0 human cDNA sequencing project";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RC TISSUE=Lymph, and Pancreas;

ALIGNMENTS

RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnarlatne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=20552140; PubMed=11103777;
 RA McConnell B.B., Vertino P.M.;
 RA "Activation of a caspase-9-mediated apoptotic pathway by subcellular
 RT redistribution of the novel caspase recruitment domain protein TRS1";
 RL Cancer Res. 60:6243-6247(2000).
 CC -!- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
 CC activity is mediated predominantly through the activation of
 CC caspase 9.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
 CC a redistribution from the cytoplasm to the aggregates occurs.
 CC These appeared as hollow, perinuclear spherical, ball-like
 CC structures.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9ULZ3-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9ULZ3-2; Sequence=VSP_004119;
 CC Name=3;
 CC IsoId=Q9ULZ3-3; Sequence=VSP_004118;
 CC Notes=Inferred from the cDNA sequence of Ref.5;
 CC -!- TISSUE SPECIFICITY: Widely expressed at low levels. Detected in
 CC peripheral blood leukocytes, lung, small intestine, spleen,
 CC thymus, colon and at lower levels in placenta, liver and kidney.
 CC Very low expression in skeletal muscle, heart and brain. Detected
 CC in the leukemia cell lines HL-60 and U937, but not in Jurkat T-
 CC cell lymphoma and Daudi Burkitt's lymphoma. Detected in the
 CC melanoma cell line WM35, but not in WM793. Not detected in HeLa
 CC cervical carcinoma cells and Molt 4 lymphocytic leukemia cells.
 CC -!- MISCELLANEOUS: In breast tumorigenesis, methylation-mediated
 CC silencing may affect genes and proteins that act as positive
 CC mediators of cell death.
 CC -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 CARD domain.
 CC -!- CAUTION: Ref.5 sequence differs from that shown due to a
 CC frameshift in position 4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB023416; BAA87339.2; -;
 DR EMBL; AF184072; AAG01187.1; -;
 DR EMBL; AF184073; AAG01188.1; -;
 DR EMBL; AF255794; AAF99665.1; -;

DR EMBL; AF310103; AAG30286.1; -;
 DR EMBL; AF384665; AAK63850.1; -;
 DR EMBL; AK000211; BAA91012.1; ALT_FRAME.
 DR EMBL; BC004470; AAH04470.1; -;
 DR EMBL; BC013569; AAH13569.1; ALT_INIT.
 DR MIM; 606838; -;
 DR GO; GO:0006917; P:induction of apoptosis; TAS.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR004020; PAAD DAPIN_dom.
 DR Pfam; PF02758; PAAD DAPIN; 1.
 DR PROSITE; PS02009; CARD; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 KW Apoptosis; Anti-oncogene; Alternative splicing.
 FT DOMAIN 1 91
 FT DAPIN.
 FT VARSPPLIC 107 195
 FT CARD.
 FT VARSPPLIC 26 85
 FT Missing (in isoform 3).
 FT FTID=VSP_004118.
 FT VARSPPLIC 93 111
 FT Missing (in isoform 2).
 FT FTID=VSP_004119.
 FT
 SQ SEQUENCE 195 AA; 21627 MW; 455987286586F46A CRC64;
 Query Match 100.0%; Score 990; DB 1; Length 195;
 Best Local Similarity 100.0%; Pred. No. 4.8e-79;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDALDALENTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSPY 60
 DB 1 MGRARDALDALENTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSPY 60
 QY 61 LETYGAELTANYLRDMGLQEMAGQIQAATHQSGAAGAPAGIQAPPOSAAPKGLHFDQHRA 120
 DB 61 LETYGAELTANYLRDMGLQEMAGQIQAATHQSGAAGAPAGIQAPPOSAAPKGLHFDQHRA 120
 QY 121 ALIARVTNVWLLDALYKGLVTDEQYQAVRAEPTNPKMKLFSPTPAWNTCKDLLOA 180
 DB 121 ALIARVTNVWLLDALYKGLVTDEQYQAVRAEPTNPKMKLFSPTPAWNTCKDLLOA 180
 QY 181 LRESOSYLVEDLERS 195
 DB 181 LRESOSYLVEDLERS 195
 RESULT 2
 ASC_MOUSE
 ID ASC_MOUSE STANDARD; PRT; 193 AA.
 AC Q9EFB4; Q9D2W9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Apoptosis-associated speck-like protein containing a CARD (mASC)
 DE (PYCARD).
 GN ASC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C; TISSUE=Thymus;
 RX MEDLINE=20580347; PubMed=11139337;
 RA Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;
 RT "Murine ortholog of ASC, a CARD-containing protein, self-associates
 RT and exhibits restricted distribution in developing mouse embryos.";
 RL Exp. Cell Res. 262:128-133(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Breast tumor;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "Pycard a PYD and CARD containing molecule";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas, and Tongue;

MEDLINE=21095660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga C., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [4]
RN SSEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RN human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
CC activity is mediated predominantly through the activation of
CC caspase 9 (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
CC a redistribution from the cytoplasm to the aggregates occurs.
CC These appeared as hollow, perinuclear spherical, ball-like
CC structures (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in small intestine, colon, thymus,
CC spleen, brain, heart, skeletal muscle, kidney, lung and liver.
CC -1- DEVELOPMENTAL STAGE: Strongly expressed at E9.5 day in the
CC telencephalon, thalamic areas of the diencephalon, heart and
CC liver.
CC -1- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
CC -1- SIMILARITY: Contains 1 DAPIN domain.
CC -1- SIMILARITY: Contains 1 CARD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB032249; BAB16609.1; -;
DR EMBL; AF301004; AAG30287.1; -;
DR EMBL; AK009852; BAB26543.1; -;
DR EMBL; AK007742; BAB25229.1; -;
DR EMBL; AK018682; BAB31341.1; -;

DR	ENBL; BCO08252; AAH08252.1; -.
DR	MGD; MGI:1931465; Asc.
DR	GO; GO:0005829; C:cytosol; IDA.
DR	InterPro; IPR001315; CARD.
DR	InterPro; IPR004020; PAAD DAPIN_dom.
DR	Pfam; PF02758; PAAD DAPIN; 1.
DR	PROSITE; PS50209; CARD; 1.
DR	PROSITE; PS50824; DAPIN; 1.
DR	Apoptosis; Anti-oncogene.
KW	DOMAIN 1 91 DAPIN.
FT	DOMAIN 105 193 CARD.
FT	CONFLICT 159 159 K -> E (IN REF. 3).
FT	SEQUENCE 193 AA; 21458 MW; 2A4EA40194870B31 CRC64;
SQ	SEQUENCE 193 AA; 21458 MW; 2A4EA40194870B31 CRC64;

Query Match 69.1%; Score 684; DB 1; Length 193;
Best Local Similarity 71.8%; Pred. No. 1.8e-52;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

Qy	1	MGARDAILDALNLTAEELKKFKLLKLVPLREGYGRIPRGALLSNDALDITDKLVSFY	60
Dd	1	MGARDAILDALNLSGDELKXPKMKLLTVQLREGYGRIPRGALLQMDAIDLTDKLVSYY	60
Qy	61	LEYTGAEILTANVLRLDMGLSQMAGLOQAATHOGSGAAPAGIQAPQSAAKPGLHFIDQHRA	120
Dd	61	LESYGUEILTWVLRDLMLGELAQQLQ-TTKEESGAVAAAASVPAQSARTGTG-HFVDQHRQ	118
Qy	121	ALIARVTNVLEWLDALYKGLVTDEQYQAVRAEPTNPSPKMKLFSTPAWNWTCKDLLQA	180
Dd	119	ALIARVTEVDGVLDALHGSLTEGQYQAVRAETTSQDKMKLFSFVPWNLTCKDSLLQA	178
Qy	181	LRESQSYLVEDLERS 195	
Dd	179	LKEIHPYLVMDEQS 193	

RESULT 3

NALI_HUMAN	STANDARD; PRT; 1473 AA.
ID	NALI_HUMAN STANDARD; Q9BZZ8; Q9BZZ9; Q9HAV8; Q9UFT4; Q9Y2E0;
AC	Q9C000; Q9BZZ8; Q9BZZ9; Q9HAV8; Q9UFT4; Q9Y2E0;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-forming cdd-4-like apoptosis protein) (Nucleotide-binding domain and caspase recruitment domain) (Caspase recruitment domain protein 7).
GN	NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_Taxid=9606;
RX	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 2).
EX	MEDLINE=21169419; PubMed=11270363;
RA	Bertin J., DiStefano P.S.;
RT	"The PYRIN domain: a novel motif found in apoptosis and inflammation proteins.";
RL	Cell Death Differ. 7:1273-1274(2000).
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORM 2).
RX	MEDLINE=21148093; PubMed=11250163;
RA	Martinon F., Hofmann K., Tschoopp J.;
RT	"the pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation.";
RL	Curr. Biol. 11:R118-R120(2001).
RN	[3]
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC	TISSUB=Erythro leukemia;
RX	MEDLINE=21153743; PubMed=11076957;
RA	Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A., Shi M.M., Vincenz C., Ward P.A.;
RT	"Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis proteins.";
RL	

J. Biol. Chem. 276:9230-9238(2001).

[4] SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.

TISSUE=P-cell;

RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S., Godzik A., Reed J.C.;

RA "A novel enhancer of the Apaf1 apoptosome involved in cytochrome c-dependent caspase activation and apoptosis."

RT J. Biol. Chem. 276:9239-9245(2001).

[5] SEQUENCE FROM N.A. (ISOFORM 2).

TISSUE=Brain;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RA "Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RT DNA Res. 6:63-70(1999).

[6] SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).

TISSUE=Uterus;

RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;

RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Able to form cytoplasmic structures termed death effector filaments. Enhances APAF1 and cytochrome c-dependent activation of pro-caspase-9 and consecutive apoptosis. Seems to bind ATP.

CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9 and with APAF1 in a cytochrome c-inducible way leading to the formation of an apoptosome. This interaction may be ATP-dependent.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Name=1; Synonyms=NAC beta, DEFCAP-L;

CC IsoId=Q9C000-1; Sequence=Displayed;

CC Name=2; Synonyms=NAC alpha, DEFCAP-S;

CC IsoId=Q9C000-2; Sequence=VSP_004327;

CC Name=3; Synonyms=NAC gamma;

CC IsoId=Q9C000-3; Sequence=VSP_004326;

CC Name=4; Synonyms=NAC delta;

CC IsoId=Q9C000-4; Sequence=VSP_004326;

CC -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are expressed in peripheral blood leukocytes, chronic myelogenous leukemia cell line K-562, followed by thymus, spleen and heart. Also detected in lung, placenta, small intestine, colon, kidney, liver and muscle.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 1 CARD domain.

CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF298548; AAG15254.1; -

DR EMBL; AF310105; AAG30288.1; -

DR EMBL; AF229059; AAK00748.1; -

DR EMBL; AF229060; AAK00749.1; -

DR EMBL; AF229061; AAK00750.1; -

DR EMBL; AF229062; AAK00751.1; -

DR EMBL; AB023143; BAA76770.1; -

DR EMBL; AL117470; CAB55945.1; -

DR FIR; T17255; T17255.

DR HSP; P13489; I44Y.

DR MIM; 606636; -

GO; GO:0005622; C:intracellular; IC.

GO; GO:0016506; F:apoptosis activator activity; NAS.

GO; GO:0008656; F:caspase activator activity; NAS.

GO; GO:0019899; F:enzyme binding activity; IPI.

GO; GO:0006919; F:caspase activation; NAS.

GO; GO:0006917; F:induction of apoptosis; NAS.

InterPro; IPR001315; CARD.

InterPro; IPR000767; Disease_resist.

InterPro; IPR001611; LRR.

InterPro; IPR007091; LRR_RNinh.

InterPro; IPR007111; NACHT_NTPase.

InterPro; IPR004020; PAAD_DAPIN_dom.

Pfam; PF00560; LRR; 2.

Pfam; PF02758; PAAD_DAPIN; 1.

PRINTS; PR00364; DISEASERSIST.

PROSITE; PS02029; CARD; 1.

PROSITE; PS50824; DAPIN; 1.

PROSITE; PS50837; NACHT; 1.

Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;

Alternative splicing.

DOMAIN 1 92 DAPIN.

DOMAIN 328 637 NACHT.

REPEAT 704 725 LRR 1.

REPEAT 807 830 LRR 2.

REPEAT 864 887 LRR 3.

REPEAT 921 944 LRR 4.

REPEAT 950 973 LRR 5.

REPEAT 1159 1215 LRR 6.

REPEAT 1216 1236 LRR 7.

DOMAIN 1374 1463 CARD.

NP_BIND 334 341 ATP (POTENTIAL).

VARSPLIC 958 987 Missing (in isoform 3 and isoform 4).

VARSPLIC 1262 1305 /FTID=VSP_004326.

VARSPLIC 340 340 Missing (in isoform 2 and isoform 3).

MUTAGEN 340 340 /FTID=VSP_004327.

MUTAGEN 340 340 K->L: NO EFFECT.

CONFLICT 155 155 K->S: NO EFFECT.

CONFLICT 246 246 L->H (IN REF. 1).

CONFLICT 782 782 T->S (IN REF. 1).

CONFLICT 878 878 T->S (IN REF. 1).

CONFLICT 995 995 T->M (IN REF. 1).

CONFLICT 1119 1119 T->I (IN REF. 1).

CONFLICT 1184 1184 M->V (IN REF. 1).

CONFLICT 1241 1241 M->V (IN REF. 1 AND 6).

CONFLICT 1366 1366 V->L (IN REF. 1).

CONFLICT 1473 1473 R->C (IN REF. 1).

SEQUENCE 1473 AA; 165865 MW; 438F0DCE45C2562D CRC64;

Query Match 25.3%; Score 250; DB 1; Length 1473;

Best Local Similarity 44.5%; Pred. No. 9.6e-14;

Matches 65; Conservative 23; Mismatches 48; Indels 10; Gaps 4;

QY 54 DKLVSVFVETVYGAELTANV--LRDMGLQEMA----GOLQATHQSGNAPAGIOAPPQS 106.

DB 1317 DOLFSEFVGHVGLSGIRLQVKDKKDETLVWEALVKGDLMPAT---TIPPARIAVPSPL 1373

QY 107 AAKPGHFDHRAALIAVNTNVEMLLDALYGVKLTDEQYQAVRAEPTNPSQMKLFSFT 166

DB 1374 DAPQLHFDVQREGIARVTSVEVLDKLGQVLSEQYERVLAEPTPSQMKLFSLS 1433

QY 167 PAWNVTCKDLLLQALRESQSYLVEDL 192

DB 1434 QSMDRKCKDGLYQALKETHPHLIMEL 1459

RESULT 4

ASC BRARE

ID ASC BRARE STANDARD; PRT; 203 AA.

Q915N6

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Apoptosis-associated speck-like protein containing a CARD (PYCARD).

GN ASC OR ASC1.


```
Matches 32; Conservative 17; Mismatches 25; Indels 4; Gaps 2;
QY 4 ARDAILDALNTAEELKFKLLSVPLREGYGR-IPRGALLSMDALDLDTKLVSYFYLE 62
DB 15 ARELLAALQDLSEQQLKFRHKLRLDAPLD---GRSIPWGLRSHSDAVALDLDTKLVFYAP 71
QY 63 TYGAELTANVLRLDMGLQE 80
DB 72 EPAVDVTRKLLKADIRD 89

RESULT 6
PYA5 MOUSE
ID PYA5 MOUSE STANDARD; PRT; 843 AA.
AC Q91W52; Q8K0L4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 5-like.
GN PYPAP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP CONCEPTUAL TRANSLATION OF 1-287.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RP SEQUENCE OF 288-843 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAP5.
RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT "Identification of mammalian orthologs associates PYPAP5 with distinct
functional roles";
RL FEBS Lett. 538:173-177(2003).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B (By similarity).
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
sequences, in analogy to ortholog sequences.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; BC013519; AAH13519.1; -;
DR EMBL; BC011139; AAH11139.1; ALT_INIT.
DR MGI; MGI:2141990; Pypaf5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 11 102 DAPIN.
FT DOMAIN 168 485 NACHT.
FT DOMAIN 569 585 ASP/GLU-RICH.
FT DOMAIN 654 661 POLY-LYS.
FT REPEAT 434 459 LRR 1.
FT REPEAT 609 632 LRR 2.
FT REPEAT 811 834 LRR 3.
FT NP_BIND 174 181 ATP (POTENTIAL).
SQ SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;

Query Match 13.3%; Score 132; DB 1; Length 843;
Best Local Similarity 41.0%; Pred. No. 0.00089;
Matches 32; Conservative 14; Mismatches 28; Indels 4; Gaps 2;

QY 4 ARDAILDALNTAEELKFKLLSVPLREGYGR-IPRGALLSMDALDLDTKLVSYFYLE 62
DB 15 ARELLAALQDLSEQQLKFRHKLRLDAPLD---GRSIPWGLRSHSDAVALDLDTKLVFYAP 71
QY 63 TYGAELTANVLRLDMGLQE 80
DB 72 VPAVEMTRQLKRSIRD 89

RESULT 7
PYA5 HUMAN
ID PYA5 HUMAN STANDARD; PRT; 892 AA.
AC P59044;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 5.
GN PYPAP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT "PYPAP7, a novel PYRIN-containing Apaf1-like protein that regulates
activation of NF-kappa B and caspase-1-dependent cytokine
processing";
RL J. Biol. Chem. 277:29874-29880(2002).
RN [2]
RP FUNCTION.
RX MEDLINE=22275822; PubMed=12387869;
RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA DiStefano P.S., Bertin J.;
RT "Functional screening of five PYPAP family members identifies PYPAP5
as a novel regulator of NF-kappa B and caspase-1.";
RL FEBS Lett. 530:73-78(2002).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
much lower levels in T-cells.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
```

Buckingham J.M., Moyzis R.K., Deaven L.L., Doggett N.A.;
"Ancient missense mutations in a new member of the Rokit gene family
are likely to cause familial Mediterranean fever";
Cell 90:797-807(1997).
[2]
SEQUENCE OF 305-754 FROM N.A., AND VARIANTS FMF.
MEDLINE=97434208; PubMed=9288094;
Bernot A., Clepet C., Dasilva C., Devaud C., Petit J.-L.,
Caloustian C., Cruaud C., Samson D., Pulcin F., Weissenbach J.,
Heilig R., Notanicola C., Domingo C., Rozenbaum M., Bencherit E.,
Topaloglu R., Dewalle M., Dross C., Hadjari P., Dupont M.,
Demaillie J., Toutou J., Smouli N., Nedelec B., Mery J.-P.,
Chabouni H., Delpech M., Grateau G.;
"A candidate gene for familial Mediterranean fever";
Nat. Genet. 17:25-31(1997).
[3]
VARIANTS FMF, AND VARIANT GLN-202.
MEDLINE=98334552; PubMed=9668175;
Bernot A., da Silva C., Petit J.-L., Cruaud C., Caloustian C.,
Casert V., Ahmed-Arab M., Dross C., Dupont M., Cattani D., Smouli N.,
Dode C., Pecheux C., Nedelec B., Medaxian J., Rozenbaum M., Rosner I.,
Delpech M., Grateau G., Demaille J., Weissenbach J., Toutou J.;
"Non-founder mutations in the MFV gene establish this gene as the
cause of familial Mediterranean fever (FMF)";
Hum. Mol. Genet. 7:1317-1325(1998).
[4]
VARIANTS FMF ILE-680; ILE-681; ILE-694; VAL-694; MET-694 DEL AND
ALA-726.
MEDLINE=99149053; PubMed=10024914;
Booth D.R., Gilmore J.D., Booth S.E., Pepys M.B., Hawkins P.N.;
"Pyriminogen mutations in familial Mediterranean fever";
QJM 91:603-606(1998).
[5]
VARIANTS FMF.
MEDLINE=99123341; PubMed=10090880;
Aksentjevich I., Torosyan Y., Samuels J., Centola M., Pras E.,
Chae J.-J., Oudoux C., Wood G., Azzaro M.P., Palumbo G., Giustolisi R.,
Pras M., Ostrer H., Kastner D.L.;
"Mutation and haplotype studies of Familial Mediterranean Fever reveal
new ancestral relationships and evidence for a high carrier frequency
with reduced penetrance in the Ashkenazi Jewish population";
Am. J. Hum. Genet. 64:949-962(1999).
[6]
VARIANTS FMF GLN-148; SER-369; GLN-408; LEU-479; ILE-680; VAL-694;
ALA-726 AND HIS-761.
MEDLINE=99294585; PubMed=10364520;
Cazeneuve C., Sarkisian T., Pecheux C., Dervichian M., Nedelec B.,
Reinert P., Avayazyan A., Kouyoumdjian J.-C., Ajrapetyan H.,
Delpech M., Goossens M., Dode C., Grateau G., Amselem S.;
"MEFV-Gene analysis in Armenian patients with Familial Mediterranean
fever: diagnostic value and unfavorable renal prognosis of the M694V
homozygous genotype-genetic and therapeutic implications";
Am. J. Hum. Genet. 65:88-97(1999).
[7]
VARIANTS FMF ILE-680; ILE-694; VAL-694 AND ALA-726.
MEDLINE=99250763; PubMed=10234504;
Shohat M., Magal N., Shohat T., Chen X., Dagan T., Mimouni A.,
Danon Y., Lotan R., Ogur G., Sirin A., Schlezinger M., Halpern G.J.,
Schwabe A., Kastner D., Rotter J.I., Fischel-Ghodsian N.;
"Phenotype-genotype correlation in familial Mediterranean fever:
evidence for an association between Met694Val and amyloidosis";
Eur. J. Hum. Genet. 7:287-292(1999).
[8]
VARIANTS FMF GLN-148; ILE-680; ILE-694; ARG-695; ALA-726 AND
HIS-761.
MEDLINE=20081069; PubMed=10612841;
Akar N., Misiroglu M., Valsinkaya F., Akar E., Cakar N., Tumer N.,
Akcaus M., Tastan H., Matzner Y.;
"MEFV mutations in Turkish patients suffering from Familial
Mediterranean Fever";
Hum. Mutat. 15:118-119(2000).
[9]
VARIANT GLN-148.

RX MEDLINE=20202844; PubMed=10737995;
 RA Ben-Cherit E., Lerer I., Malamud E., Domingo C., Abeliovich D.;
 RT "The E148Q mutation in the MEV gene: is it a disease-causing mutation
 or a sequence variant?";
 RL Hum. Mutat. 15:385-386(2000).
 CC -!- FUNCTION: LIKELY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC DIFFERENTIATED GRANULOCYTES.
 CC -!- SUBCELLULAR LOCATION: Nucleus (Probable).
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PERIPHERAL BLOOD
 CC LEUCOCYTES, PARTICULARLY IN MATURE GRANULOCYTES BUT NOT IN
 CC MONOCYTES AND LYMPHOCYTES. NOT EXPRESSED IN SPLEEN, THYMUS,
 CC PROSTATE, TESTIS, OVARY, SMALL INTESTINE, COLON, HEART, BRAIN,
 CC PLACENTA, LUNG, LIVER, MUSCLE, KIDNEY, PANCREAS.
 CC -!- DISEASE: Defects in MEV are the cause of familial Mediterranean
 CC fever (FMF) [MIM:249100], an autosomal recessive inherited
 CC disorder characterized by recurrent episodic fever, serosal
 CC inflammation and pain in the abdomen, chest or joints. It is
 CC frequently complicated by amyloidosis, which leads to renal
 CC failure and can be prophylactically treated with colchicine. FMF
 CC primarily affects ethnic groups living around the Mediterranean
 CC basin: North-African Jews, Armenians, Arabs and Turks.
 CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- DATABASE: NAME=GeneDis; NOTE=FMF;
 CC WWW="http://life2.tau.ac.il/GeneDis/Tables/FMF/fmf.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF018080; AB070557.1; --
 CC EMBL: Y14441; CAAY4793.1; --
 CC EMBL: AJ003147; CAA05906.1; --
 CC EMBL: AF111163; AAD26152.1; --
 CC Genew; HGNC:6998; MEV.
 CC MIM: 249100; --
 CC GO: GO:0005875; C:microtubule associated complex; IDA.
 CC GO: GO:0005634; C:nucleus; IDA.
 CC GO: GO:0003779; F:actin binding activity; IDA.
 CC GO: GO:0008270; F:zinc ion binding activity; NAS.
 CC GO: GO:0006954; P:inflammatory response; IDA.
 CC InterPro: IPR001870; B302.
 CC InterPro: IPR004020; PAAD_DAPIN_dom.
 CC InterPro: IPR006574; PRY.
 CC InterPro: IPR003877; SPRY receptor.
 CC InterPro: IPR000315; Znf_Bbox.
 CC Pfam: PF02758; PAAD_DAPIN; 1.
 CC Pfam: PF00622; SPRY; 1.
 CC Pfam: PF00643; zf-B_box; 1.
 CC SMART: SM00336; BBOX; 1.
 CC SMART: SM00589; PRY; 1.
 CC SMART: SM00449; SPRY; 1.
 CC PROSITE: PS50824; DAPIN; 1.
 CC PROSITE: PS50119; ZF_BBOX; 1.
 KW Nuclear protein; zinc-finger; Polymorphism; Disease mutation.
 FT DOMAIN 1 92 DAPIN.
 FT ZN_FING 370 412 B BOX-TYPE.
 FT DOMAIN 420 437 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT VARIANT 148 148 E -> Q (in FMF; common mutation;
 FT associated with Ser-369 and Gln-408 in
 FT cis; could be a polymorphism).
 FT FTID=VAR_009051.
 FT E -> D (in FMF).
 FT FTID=VAR_009052.
 FT R -> Q (in FMF; common mutation;
 FT associated with Ser-369 and Gln-408 in
 FT cis; could be a polymorphism).
 FT FTID=VAR_009053.
 FT T -> I (in FMF).
 FT FTID=VAR_009054.
 FT P -> S (in FMF; reduced penetrance among

FT FTID=VAR_009055.
 FT R -> Q (in FMF; associated with Gln-148
 FT and Ser-369 in cis; could be a
 FT polymorphism).
 FT FTID=VAR_009056.
 FT F -> L (in FMF).
 FT FTID=VAR_009057.
 FT M -> I (in FMF).
 FT FTID=VAR_009058.
 FT T -> I (in FMF).
 FT FTID=VAR_009059.
 FT Missing (in FMF).
 FT FTID=VAR_009060.
 FT M -> I (in FMF).
 FT FTID=VAR_009061.
 FT M -> V (in FMF; very common mutation
 FT particularly in North-African Jews; can
 FT be associated with amyloidosis
 FT development).
 FT FTID=VAR_009062.
 FT Missing (in FMF).
 FT FTID=VAR_009063.
 FT K -> R (in FMF; reduced penetrance among
 FT Ashkenazi Jews).
 FT FTID=VAR_009064.
 FT V -> A (in FMF; common mutation; found in
 FT Iraqi and Ashkenazi Jews, Druze,
 FT Armenians).
 FT FTID=VAR_009065.
 Query Match 12.4%; Score 122.5; DB 1; Length 781;
 Best Local Similarity 30.8%; Pred. No. 0.0055;
 Matches 33; Conservative 18; Mismatches 53; Indels 3; Gaps 1;
 QY 6 DAIDLALNTAEELKFKLLSVLPREGYGRIPRGALLSMDALDLTDKLVSFLEYTG 65
 Db 7 DHLSTLEELVPYDFEKFELQNTSVQKHSRPSQIQARPVKMATLLVYYGEYA 66
 QY 66 AELTANVLRLMGLQEMACQLQAATHQSGAAGIQAAPPQSAAPGL 112
 Db 67 VQLTLQVLRAINQRLLEELHRAAIQEYSTQENGTD---DSAASSL 110
 RESULT 9
 ID PYA3 HUMAN STANDARD; PRT; 980 AA.
 AC Q8WX94;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PYRIN-containing APAF1-like protein 3.
 GN PYPAF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2216247; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing.";
 RL J. Biol. Chem. 277:29874-29880(2002).
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```

DR EMBL; AY026322; AAK08982.1; -
DR EMBL; AF405558; AAL02427.1; -
DR EMBL; AF511652; AAM46959.1; -
DR InterPro; IPR001315; CARD.
DR Pfam; PF00619; CARD; 1.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS0209; CARD; 1.
KW Apoptosis; Nuclear protein; Alternative splicing.
FT DOMAIN 340 430
FT VARSPLIC 282 286
FT VARSPLIC 287 431
FT VARSPLIC 287 431
FT MUTAGEN 366 366
FT CONFLICT 60 60
FT CONFLICT 326 326
FT CONFLICT 422 422
SQ SEQUENCE 431 AA; 48932 MW; CB54D130807732E6 CRC64;

Query Match
Best Local Similarity 24.0%; Score 99; DB 1; Length 431;
Matches 42; Conservative 32; Mismatches 73; Indels 28; Gaps 5;

QY 21 KPFKLLSVPLREGVGRIPRGALLSMDALDITDKLVSPYLETYGAEITANVLRDMGLQE 80
DB 281 KELKLSYSPGEIQHPSKFYAGQKEPIQLEITEK-----RHGTLVMDTEVKPVDLQL 333
QY 81 MACOLQAATHQSGGAAPAGIQAPPSAAKPGHLFIDHRAALIAITVNVVWLLDALY-CK 139
DB 334 VA-----ASAP-----PPFS-----GAAFVKNHRQLQARMGDLKGLVLDLQDNE 373
QY 140 VLTDEQVAVRAEPTNPSKMRKLFSTPAWNWTKDLLLQALRESOSYLVEDLER 194
DB 374 VLTENEKELVEQETKQSKNEALLSVKKGDLALDVLFSISERDPYLVSLYLRQ 428

RESULT 11
SIDE_SALT
ID -SIDE_SALT STANDARD; PRT; 143 AA.
AC Q9ZAN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Silver-binding protein silc precursor.
GN SILE
OS Salmonella typhimurium.
OG Plasmid pMG101.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]_
RP SEQUENCE FROM N.A., SEQUENCE OF 21-35, AND CHARACTERIZATION.
RX MEDLINE=99128056; PubMed=9930866;
RA Gupta A., Matsui K., Lo J.-F., Silver S.;
RT "Molecular basis for resistance to silver cations in Salmonella.";
RL Nat. Med. 5:183-188(1999).
CC -!- FUNCTION: COMPONENT OF THE SIL CATION-EFFLUX SYSTEM THAT CONFERS
CC RESISTANCE TO SILVER.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- INDUCTION: By silver.
CC -!- SIMILARITY: TO E.COLI PCOE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF067954; AAD1743.1; -
DR EMBL; AF067954; AAD1743.1; -
KW Metal-binding; Periplasmic; Signal; Plasmid.
FT SIGNAL 1 20

```

```

FT CHAIN 21 143 SILVER-BINDING PROTEIN SILE.
SQ SEQUENCE 143 AA; 15201 MW; 703B9EC0F8BDOCB6 CRC64;

Query Match
Best Local Similarity 9.2%; Score 91.5; DB 1; Length 143;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YCAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPSAAKPGHLFIDHRAAL 122
DB 12 GLISSAWATETVNIHERVNNAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66
QY 123 IARVTNVWLLDALYGVKVLTDQYQ--AVRAEPTNPSK 158
DB 67 IAHETMTNGSADA-HQKMWESHQRMGMSQTVSPGPSK 103

RESULT 12
RUVB_THETN
ID RUVB_THETN STANDARD; PRT; 338 AA.
AC Q8RAN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase RUVB.
GN RUVB OR TTE1180.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: The ruva-ruvb complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RUVAB is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -!- SUBUNIT: Forms a complex with ruva (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RUVB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE013080; AAM24411.1; -
DR HAMAP; MF 00016; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004605; RUVB.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00635; ruvb; 1.
KW DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
FT NP_BIND 59 66 ATP (POTENTIAL).
FT BIND 59 66
SQ SEQUENCE 338 AA; 38072 MW; 2B9562D8CDB988D1 CRC64;

Query Match
Best Local Similarity 9.0%; Score 89; DB 1; Length 338;
Matches 39; Conservative 32; Mismatches 73; Indels 18; Gaps 5;

QY 3 RARDAILDALNTAEELK---KFKLLSVPLREG-----YGRIPRGALLSMDAL----- 50

```

```
Db 169 RSRFGVNRDYSVEELKEIIRKSANILNIGIDDAEFAIARRSGTGPRIANRLKVR 228
Qy 51 DLTDKLVSYLYTYGAEITANVL--RDMGLQMGLOQAATHQGGGAAPAGIQAPPQSA 108
Db 229 DPAEVKNGYIDYNTANTALNMLGVDEMDIKILIAIEKEGGGPGVGDIAASVG 288
Qy 109 KPELHFIDQRAALIRVNTNVEWLLDALYKVLTDQYQAVR 150
Db 289 EBDGTIEDMYEPLM----QIGFLNRPGRVVTKLAYQYLK 326

RESULT 13
M22_STRPY STANDARD; PRT; 372 AA.
AC P50469;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE M protein, serotype 2.2 precursor.
GN EML2.2.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T2/44/RB4/119;
RA MEDLINE=92104662; PubMed=1370269;
RA Bessen D.E., Fischetti V.A.;
RT "Nucleotide sequences of two adjacent M or M-like protein genes of group A streptococci: different RNA transcript levels and identification of a unique immunoglobulin A-binding protein.";
RL Infect. Immun. 60:124-135(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Fischetti V.A., Bessen D.E.;
RT "Immunoglobulin A binding protein.";
RL Patent number US5556944, 17-SEP-1996.
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO PHAGOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
CC -!- SIMILARITY: TO OTHER M PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61276; CAA43582.1; -
DR EMBL; I26204; -; NOT_ANNOTATED_CDS.
DR PIR; S23326; S23326.
DR InterPro; IPR005877; Gpos_YsIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR InterPro; IPR003345; M_repeat.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02370; M; 8.
DR Pfam; PF04650; YsIRK_signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRPFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRPFAMS; TIGR01168; YsIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Virulence; Phagocytosis; Cell wall; Peptidoglycan-anchor; Repeat;
KW Antigen; Coiled coil; signal.
FT SIGNAL 1 41 POTENTIAL.
FT CHAIN 42 342 M PROTEIN, SEROTYPE 2.2.
```

```
FT PROPEP 343 372 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 131 244 3 X REPEATS, TYPE C.
FT REPEAT 131 153 C-1.
FT REPEAT 173 195 C-2.
FT REPEAT 222 244 C-3.
FT DOMAIN 305 338 GLY/PRO-RICH (CELL WALL-SPANNING) (BY SIMILARITY).
FT SITE 339 343 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 342 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 372 AA; 41149 MW; E8FD5D0920C95C74 CRC64;

Query Match 8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 3.1;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

Qy 12 LENTAEELKKFKLLSVPLREGYGRIPRGALLSMDAL-----DLTDKLVSYLYETGA 66
Db 78 LEKINAEKEKKLEAINKELNENYKLDG-----IDALEKEKEDLTKTKTKEINIS 133
Qy 67 ELTANVL-RDMGLQMGLOQAATHQGGGAAPAGIQAPPQ--SAAKPGL-HFIDQRAAL 122
Db 134 EASRKGLSRDLEASRTAKKEAKHQLKLEAKKLTGEGVSEASRKGLSNDLEASRAA- 192
Qy 123 IARVTNVEWLLDALYKVLTDQ-----YQAVRAEPTNPSPKMKLFSTPANNWTCKOLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHQLKLEADYQVSETSRKGLS-----RD-- 234
Qy 178 LQALRESQSYLVEDLERS 195
Db 235 LEASREANKVTSSELTOA 252

RESULT 14
SYFB_CORGL STANDARD; PRT; 835 AA.
ID SYFB CORGL STANDARD; PRT; 835 AA.
AC Q8NQ6;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE (Phenylalanine-tRNA ligase beta chain) (PheRS).
GN PHET OR CGLI390.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP + diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain family. Subfamily 1.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005278; BAB98783.1; ALT_INIT.
DR HAMAP; MF_00283; -; 1.
DR InterPro; IPR005146; B3_4.
```

DR InterPro; IPR005147; B5.
 DR InterPro; IPR005121; Fdx-AntiCB.
 DR InterPro; IPR004532; PheT bact.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF03483; B3_4; 1.
 DR Pfam; PF03484; B5; 1.
 DR Pfam; PF03147; FDX-ACB; 1.
 DR Pfam; PF01588; tRNA_bind; 1.
 DR TIGRFAMs; TIGR00472; pHeT_bact; 2.
 DR PROSITE; PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Metal-binding; Magnesium; RNA-binding; tRNA-binding;
 KW Complete proteome.
 FT DOMAIN 44 160 TRNA-BINDING.
 FT METAL 472 472 MAGNESIUM (BY SIMILARITY)
 FT METAL 478 478 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
 SIMILARITY).
 FT METAL 481 481 MAGNESIUM (BY SIMILARITY).
 FT METAL 482 482 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 835 AA; 89435 MW; 285EC9A0673DA49F CRC64;
 Query Match 8.7%; Score 86.5; DB 1; Length 835;
 Best Local Similarity 24.5%; Pred. No. 7.9;
 Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7;
 QY 29 SVPLREGYGRIPRGALLSMDALDLDTKLVSYL---ETYGAEITANVLDRMGLOEMAGOL 85
 DB 114 AISARTYGRMGAMTCSASEGLADKQNSGITLDPSTG-FPGEDARQALGLEDVTFDV 172
 QY 86 QAATHOGGAAPAGIQAAPPQSAKPGHFDIHOHRAALIAVNTNVEVLVLLDLYGKVL--- 142
 DB 173 NVTDPGRYALSGRLTRELASAF-SLTFTD---PAIEPAVAGIEVKVPAVEGSLINVEL 227
 QY 143 DEOYQAVRAEPTNPQSKRLFSFTPAWN---WTCKDLLLQALR 182
 DB 228 REETKAIR-----FGLRKVSGIDPAESPFPMQRELMLSGOR 264
 RESULT 15
 CISL_MOUSE STANDARD; PRT; 1033 AA.
 AC Q8R4B8;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cold autoimmune syndrome 1 protein homolog (PYRN-containing
 DE APAFL-like protein 1) (Mast cell maturation inducible protein 1).
 GN C1A51 OR PYPAF1 OR MMIG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/cJ;
 RA Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
 RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
 RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
 RT Identification of inducible genes during in vitro maturation of mouse
 RT bone marrow-derived mast cells to connective tissue-type mast cells.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May function as a potential inducer of apoptosis.
 CC Interacts selectively with apoptosis-associated specklike protein
 CC containing a CARD domain (ASC). This complex may function as an
 CC upstream activator of NF-kappaB signaling (By similarity).
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF486632; AAL90874.1; --
 DR MGD; MGI:2653833; Class1.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR RNinh.
 DR InterPro; IPR003590; LRR RNinh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_RI; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW Apoptosis; Repeat; Leucine-rich repeat.
 FT DOMAIN 1 91 DAPIN.
 FT METAL 216 532 NACHT.
 FT REPEAT 737 760 LRR 1.
 FT REPEAT 794 817 LRR 2.
 FT REPEAT 851 874 LRR 3.
 FT REPEAT 880 903 LRR 4.
 FT REPEAT 908 931 LRR 5.
 FT REPEAT 937 964 LRR 6.
 FT REPEAT 965 988 LRR 7.
 SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B12117 CRC64;
 Query Match 8.7%; Score 86; DB 1; Length 1033;
 Best Local Similarity 35.6%; Pred. No. 11;
 Matches 21; Conservative 9; Mismatches 29; Indels 0; Gaps 0;
 QY 1 MGRARDALDLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSF 59
 DB 1 MTSVRCKLAQVLEDDVDLKKFKWHLDDYPEKGCIPVPRGQMEKADHLDLATLMIDF 59
 Search completed: January 29, 2004, 13:46:32
 Job time : 5.00185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 10.8067 Seconds
(without alignments)
4656.416 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTAEEL.....LLQLALRSQSYLVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_prodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	74.3	195	6 Q8HXX9	Q8hxx9 bos taurus
2	673	68.0	193	11 Q8CHK8	Q8chk8 rattus norv
3	275	27.8	89	4 Q8NFP8	Q8nfp8 homo sapien
4	267	27.0	89	4 Q8WCX3	Q8wcx3 homo sapien
5	142	14.3	750	11 Q9J25	Q9j25 rattus norv
6	142	14.3	767	11 Q9J26	Q9j26 mus musculu
7	104	10.5	404	13 Q9DDJ2	Q9ddj2 brachydanio
8	91.5	9.2	127	2 Q8VPY3	Q8vpy3 serratia ma
9	91.5	9.2	127	2 Q8VPW8	Q8vpw8 salmonella
10	91.5	9.2	127	2 Q8VPX8	Q8vpw8 salmonella
11	91.5	9.2	127	2 Q8VPY8	Q8vpy8 serratia ma
12	91.5	9.2	383	13 Q919L7	Q919l7 brachydanio
13	86.5	8.7	439	16 P74033	P74033 synecocyst
14	86.5	8.7	439	16 P8QNC6	P8qnc6 corynebacte
15	86	8.7	580	17 Q8TSR7	Q8tsr7 methanosarc
16	85	8.7	1033	11 Q8R4B8	Q8r4b8 mus musculu

17	85	8.6	1488	16	Q8Z7Z5	Q8z7z5 salmonella
18	85	8.6	1488	16	Q935S7	Q935s7 salmonella
19	84.5	8.5	894	10	Q8GSP4	Q8gsf4 oryza sativ
20	84.5	8.5	1241	3	Q9HEM2	Q9hem2 neurospora
21	84	8.5	292	16	Q25976	Q25976 helicobacte
22	84	8.5	530	5	Q8T9T6	Q8t9t6 aedes aegyp
23	83.5	8.4	542	17	Q9HRG1	Q9hrg1 halobacteri
24	83.5	8.4	2556	12	Q9QH56	Q9qhs6 gallid herp
25	83	8.4	469	16	Q8CVN9	Q8cvn9 escherichia
26	83	8.4	545	16	Q9CBL2	Q9cbl2 mycobacteri
27	83	8.4	877	2	Q45095	Q45095 bacillus ci
28	82.5	8.3	409	5	Q95984	Q95984 stylonychia
29	82.5	8.3	522	2	Q85210	Q85210 vibrio chol
30	82.5	8.3	522	16	Q9K394	Q9k394 vibrio chol
31	82.5	8.3	567	3	Q8NIK9	Q8nik9 cryptococcu
32	82	8.3	292	16	Q9ZJ11	Q9zj11 helicobacte
33	81	8.2	456	4	Q9BUC6	Q9buc6 homo sapien
34	81	8.2	503	2	Q9XEM8	Q9xbw8 rhodobacter
35	81	8.2	631	4	Q8NDM4	Q8ndm4 homo sapien
36	81	8.2	631	4	Q969Z0	Q969z0 homo sapien
37	81	8.2	792	16	Q8ZPS3	Q8zps3 salmonella
38	81	8.2	792	16	Q8Z6J0	Q8z6j0 salmonella
39	81	8.2	896	2	Q9AN79	Q9an79 bradyrhizob
40	81	8.2	9477	2	Q9L4X3	Q9l4x3 streptomyc
41	80.5	8.1	266	10	Q8LM81	Q8lm81 oryza sativ
42	80.5	8.1	718	16	Q9JYB6	Q9jyb6 neisseria m
43	80.5	8.1	725	16	Q9JYB6	Q9jyb6 neisseria m
44	80.5	8.1	727	16	Q8DE24	Q8de24 vibrio vuln
45	80.5	8.1	972	4	Q8N612	Q8n612 homo sapien

ALIGNMENTS

RESULT 1

ID	Q8HXX9	PRELIMINARY;	PRT;	195 AA.
AC	Q8HXX9;			
DT	01-MAR-2003 (TREMREL. 23, Created)			
DT	01-MAR-2003 (TREMREL. 23, Last sequence update)			
DT	01-MAR-2003 (TREMREL. 23, Last annotation update)			
DE	Apoptosis-associated speck-like protein containing a CARD.			
GN	BASC.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Masumoto J., Zhou W., Chen P.F., Su F., Kuwada J.Y., Hidaka E.,			
RA	Katsuyama T., Sagara J., Taniguchi S., Taniguchi S., Ngo-Hazelett P.,			
RA	Postlethwait J.H., Nunez G., Inohara N.,			
RT	"Casp-1: A zebrafish caspase activated by ASC oligomerization required			
RT	for pharyngeal Arch development."			
RL	J. Biol. Chem. 0:0-0(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20036508; PubMed=10567338;			
RA	Masumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T.,			
RA	Niikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.;			
RT	"ASC, a novel 22-kDa protein, aggregates during apoptosis of human			
RT	promyelocytic leukemia HL-60 cells,"			
RL	J. Biol. Chem. 274:33835-33838(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20580347; PubMed=11139337;			
RA	Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;			
RT	"Murine ortholog of ASC, a CARD-containing protein, self-associates,			
RT	and exhibits restricted distribution in developing mouse embryos.";			
RL	Exp. Cell Res. 262:128-133(2001).			
DR	EMBL; AB050006; BAC43753.1; -.			
SQ	SEQUENCE 195 AA; 21917 MW; 7C9D4BD8BA9A9E8 CRC64;			

```
Query Match 74.3%; Score 736; DB 6; Length 195;
Best Local Similarity 75.4%; Pred. No. 1.7e-59;
Matches 147; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRTRDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
[1]
QY 61 LETYGAEALTANVLDRMGLOEMAGOLQAATHQSGAPAGIQAPQSAAPGLHFDIQHRA 120
DB 61 LEAYGAELTALVLRDMGMEVAEQLETWMSKGRPNVLAERDPLQKTAKPGLHFVDQHRA 120
[1]
QY 121 ALIARVTNVEMLLDALYGVKLVLTDEQYQAVRAEPTNPDKMLFSPFPANNWTCOLLLOA 180
DB 121 ALIARVTNVGVLDALYGVKLVLEEQYQAVRAETSSDKMKLFSPFPANNWTCOLLLOA 180
[1]
QY 181 LRESQSYLVEDLERS 195
DB 181 LRDTQPYLVDDLEQS 195
[1]

RESULT 2
Q8CHK8 PRELIMINARY; PRT; 193 AA.
AC Q8CHK8;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Apoptosis-associated speck-like protein.
RASC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA Masumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,
RA Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazelett P.,
RA Postlethwait J.H., Nunez G., Inohara N.,
RT "Casp9: A zebrafish caspase activated by ASC oligomerization required
RT for pharyngeal Arch development.";
RL J. Biol. Chem. 0:0(2002).
DR EMBL; AB053165; BAC43754.1; -.
SQ SEQUENCE 193 AA; 21654 MW; F3B27B560D86A17B CRC64;

Query Match 68.0%; Score 673; DB 11; Length 193;
Best Local Similarity 69.7%; Pred. No. 9.7e-54;
Matches 136; Conservative 20; Mismatches 37; Indels 2; Gaps 2;

QY 1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
[1]
QY 61 LETYGAEALTANVLDRMGLOEMAGOLQAATHQSGAPAGIQAPQSAAPGLHFDIQHRA 120
DB 61 LRGYGLTSLVLRDMGLOEMAGOLQAATHQSGAPAGIQAPQSAAPGLHFDIQHRA 118
[1]
QY 121 ALIARVTNVEMLLDALYGVKLVLTDEQYQAVRAEPTNPDKMLFSPFPANNWTCOLLLOA 180
DB 119 ALIARVTNVGVLDALYGVKLVLEEQYQAVRAETPNQNMKLFSPFPANNWTCOLLLOA 178
[1]
QY 181 LRESQSYLVEDLERS 195
DB 179 LRQTPYLVTDLEQS 193
[1]

RESULT 3
Q8NFP8 PRELIMINARY; PRT; 89 AA.
AC Q8NFP8;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)

Query Match 27.0%; Score 267; DB 4; Length 89;
Best Local Similarity 64.8%; Pred. No. 5.6e-17;
Matches 57; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTKEAILKVLNLTPEELKKFKMLGTVPPLREGFGRIPRGALQDIDVLTDKLVASY 60
[1]
QY 61 LETYGAEALTANVLDRMGLOEMAGOLQA 88
DB 61 YEDYAAELVAVLRDMRMLEEAARLQRA 88
[1]

RESULT 4
Q8WXC3 PRELIMINARY; PRT; 89 AA.
AC Q8WXC3;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Pyrin-only protein 1.
GN POP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Bertin J.;
RT "POP1: a pyrin-only protein that regulates inflammatory signaling.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF454669; AAL58439.1; -.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
SQ SEQUENCE 89 AA; 10107 MW; 4CDF6D672DDDD98E CRC64;

Query Match 27.8%; Score 275; DB 4; Length 89;
Best Local Similarity 65.9%; Pred. No. 1e-17;
Matches 58; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

QY 1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTKEAILKVLNLTPEELKKFKMLGTVPPLREGFGRIPRGALQDIDVLTDKLVASY 60
[1]
QY 61 LETYGAEALTANVLDRMGLOEMAGOLQA 88
DB 61 YEDYAAELVAVLRDMRMLEEAARLQRA 88
[1]

RESULT 5
Q8JUJ25 PRELIMINARY; PRT; 750 AA.
AC Q8JUJ25;
DT 01-OCT-2000 (TremBLrel. 15, Created)
```

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 GN Pyrin (Marenostrein).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=20279845; PubMed=10818206;
 RA Chae J.J., Centola M., Aksentjevich I., Dutra A., Tran M., Wood G.,
 RA Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
 RT "Isolation, genomic organization, and expression analysis of the mouse
 RT and rat homologs of MEV, the gene for familial Mediterranean fever.";
 RL Mamm. Genome 11:428-435(2000).
 CC -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
 CC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
 CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
 CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN AND, TO A LESSER DEGREE IN
 CC THE LUNG. NOT EXPRESSED IN THYMUS, TESTIS, OVARY, HEART, BRAIN,
 CC LIVER, KIDNEY AND MUSCLE.
 CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 DR EMBL: AF143410; AAF03767.1; -
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF02758; PAAD_DAPIN; 1.
 DR SMART: SM00643; zf-B_box; 1.
 DR PROSITE: PS00824; DAPIN; 1.
 DR PROSITE: PS50119; ZF_BOX; 1.
 KW Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
 KW Microtubules; Zinc-finger; Zinc.
 FT DOMAIN 1 92 DAPIN.
 FT ZN FING 440 482 B BOX-TYPE.
 FT ZN FING 440 482 B BOX-TYPE.
 SQ SEQUENCE 750 AA; 83994 MW; C76D0F3E02711312 CRC64;
 Query Match 14.3%; Score 142; DB 11; Length 750;
 Best Local Similarity 33.7%; Pred. No. 0.00029;
 Matches 35; Conservative 19; Mismatches 50; Indels 0; Gaps 0;
 QY 6 DAILDALENLTAEELKKKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSYFLEYTG 65
 DB 7 DHLNLTLELLPYELKFKFKLHTTSLEKSHSIPLSLVKMARPIKLTLLTYGEEYA 66
 QY 66 AELTANVLDMGLQEMAGLOQAATHGSGAAPAGIQAPPOSAK 109
 DB 67 VRLTLQLRATNQRQALAEELKATGHEHLTEENGSGVQSSAE 110
 RESULT 6
 QYJ26 PRELIMINARY; PRT; 767 AA.
 AC QYJ26
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Pyrin (Marenostrein).
 GN MEV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=129/Sv;
 RX MEDLINE=20279845; PubMed=10818206;
 RA Chae J.J., Centola M., Aksentjevich I., Dutra A., Tran M., Wood G.,

RA Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
 RT "Isolation, genomic organization, and expression analysis of the mouse
 RT and rat homologs of MEV, the gene for familial Mediterranean fever.";
 RL Mamm. Genome 11:428-435(2000).
 CC -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
 CC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
 CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
 CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN PERIPHERAL BLOOD
 CC GRANULOCYTES. NOT EXPRESSED IN LYMPHOCYTES, THYMUS, TESTIS, OVARY,
 CC HEART, BRAIN, LUNG, LIVER, KIDNEY AND MUSCLE.
 CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 DR EMBL: AF143409; AAF03766.1; -
 DR MGD: MGI:1859396; Mefv.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR InterPro: IPR000315; Znf_Box.
 DR Pfam: PF02758; PAAD_DAPIN; 1.
 DR Pfam: PF00643; zf-B_box; 1.
 DR SMART: SM00336; BBOX; 1.
 DR PROSITE: PS50824; DAPIN; 1.
 DR PROSITE: PS50119; ZF_BOX; 1.
 KW Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
 KW Microtubules; Zinc-finger; Zinc.
 FT DOMAIN 1 92 DAPIN.
 FT ZN FING 439 481 B BOX-TYPE.
 FT ZN FING 439 481 B BOX-TYPE.
 SQ SEQUENCE 767 AA; 86264 MW; FF102CB3FD7C1EB CRC64;
 Query Match 14.3%; Score 142; DB 11; Length 767;
 Best Local Similarity 28.3%; Pred. No. 0.0003;
 Matches 32; Conservative 26; Mismatches 47; Indels 8; Gaps 1;
 QY 6 DAILDALENLTAEELKKKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSYFLEYTG 65
 DB 7 DHLNLTLELLPYELKFKFKLHTTSLEKSHSIPLSLVKMARPIKLTLLTYGEEYA 66
 QY 66 AELTANVLDMGLQEMAGLOQAATHGSGAAPAGIQAPPOSAK 110
 DB 67 VRLTLQLRATNQRQALAEELKATGHEHLTEENGSGVQSSVSKAKSVKVP 119
 RESULT 7
 QYJ26 PRELIMINARY; PRT; 404 AA.
 AC QYJ26
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Casp2.
 GN CASP2 OR CASPY2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20373792; PubMed=10917738;
 RA Inohara N., Nunez G.;
 RT "Genes with homology to mammalian apoptosis regulators identified in
 RT zebrafish.";
 RL Cell Death Differ. 7:509-510(2000).
 DR EMBL: AF327410; AAG45230.1; -
 DR HSSP: P29466; ICE.
 DR ZFIN: ZDB-GENE-020812-1; caspb.
 DR InterPro: IPR002398; ICE.
 DR InterPro: IPR002138; ICE_p10.
 DR InterPro: IPR001309; ICE_p20.
 DR InterPro: IPR004020; PAAD_DAPIN_dom.
 DR Pfam: PF00655; ICE_p10; 1.
 DR Pfam: PF00656; ICE_p20; 1.

```
DR Pfam; PF02758; PAAD DAPIN; 1.
DR PRINTS; PR00376; IL1BCENZYME.
DR SMART; SM00115; CASC; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01201; CASPASE_HIS; 1.
DR PROSITE; PS02071; CASPASE_P10; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
DR PROSITE; PS0824; DAPIN; 1.
DR SEQUENCE 404 AA; 46077 MW; 629BECFEB4A87D3B CRC64;

Query Match      10.5%; Score 104; DB 13; Length 404;
Best Local Similarity 30.8%; Pred. No. 0.38;
Matches 36; Conservative 17; Mismatches 50; Indels 14; Gaps 4;

QY 10 DALENITAEELKFKLLSVPLREGYGRIPGALLSMDALDITDKLVSYFLETYGAEIT 69
DB 10 DVLEDDVEAEKQF--TRQLWIGVKGVPIPRKLENKQDQVVDVSMVOQYSEDAGT-IT 67
QY 70 ANVLRDMGLQEMAGQL-----QAATHQSGAAPAGIQAPPQSAAKPGLHFDIQ 117
DB 68 VQTLRKIKQNERAKRLESNLLKVQSQQENKQNSEP---QPIPIISQPIQIIISQ 121

RESULT 8
Q8VPY3 PRELIMINARY; PRT; 127 AA.
AC Q8VPY3
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE
OS Serratia marcescens.
OG Bacteri; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in Inch incompatibility group plasmids."
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009377; AAL68934.1; -.
KW Plasmid.
FT NON TER 127
SQ SEQUENCE 127 AA; 13438 MW; 67B3822C25BFC9A5 CRC64;

Query Match      9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ--GSGAAPAGIQAPPQSAAKPGLHFDIQHRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEQYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 9
Q8VPW8 PRELIMINARY; PRT; 127 AA.
AC Q8VPW8
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE
OS Salmonella oranienberg.
OG Plasmid MIP235.

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ--GSGAAPAGIQAPPQSAAKPGLHFDIQHRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEQYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 10
Q8VPX8 PRELIMINARY; PRT; 127 AA.
AC Q8VPX8
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE
OS Salmonella enterica subsp. enterica serovar Ohio.
OG Plasmid pMIP233.
OC Bacteri; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=117541;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in Inch incompatibility group plasmids."
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009382; AAL68941.1; -.
KW Plasmid.
FT NON TER 127
SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;

Query Match      9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ--GSGAAPAGIQAPPQSAAKPGLHFDIQHRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEQYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 11
Q8VPY8 PRELIMINARY; PRT; 127 AA.
AC Q8VPY8
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
OG Bacteri; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=28147;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in Inch incompatibility group plasmids."
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009392; AAL68944.1; -.
KW Plasmid.
FT NON TER 127
SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;

Query Match      9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ--GSGAAPAGIQAPPQSAAKPGLHFDIQHRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEQYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103
```



```

DT 01-WAR-2002 (TReMBLrel. 20, Last annotation update)
DE Sile (Fragment).
GN SILE.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in inch incompatibility group
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY003372; AAL68931.1; -.
KW Plasmid.
FT NON_TER.
SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;

Query Match 9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQIQAAATHQ-GSGAAPAGIQAPPSAAKPGHFIIDHRAAL 122
Db FGLISSAWATETVTHIRVNNVQAAPAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVKVLTDQYQ--AVRAEPTNPSK 158
Db IAHETMTNGSADA-HQKMWESHQRMMGMSQTVSPTGFSK 103

RESULT 12
QY1917 PRELIMINARY; PRT; 383 AA.
AC QY1917;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Caspase.
GN CASPA OR CASPY.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish."
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF233434; AAF66964.1; -.
DR HSP; P29456; 1ICE.
DR ZFIN; ZDB-GENE-000616-3; caspa.
DR InterPro; IPR002398; ICE.
DR InterPro; IPR002138; ICE_P10.
DR InterPro; IPR001309; ICE_P20.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00655; ICE_P10; 1_.
DR Pfam; PF00656; ICE_P20; 1_.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PR00376; ILIICENZYM.
DR SMART; SM00115; CASC; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS0207; CASPASE_P10; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
DR PROSITE; PS0824; DAPIN; 1.
SQ SEQUENCE 383 AA; 43966 MW; 21890871309774C3 CRC64;

Query Match 9.2%; Score 91.5; DB 13; Length 383;
Best Local Similarity 26.2%; Pred. No. 5;
Matches 28; Conservative 24; Mismatches 44; Indels 11; Gaps 4;

QY 5 RDAILDALENITABELKKFKLLSVPLREGVGRIPRGALLSM-DALDLTDKLVSFYLET 63
Db KDLQDALSNIGADNLRFRQRLGD---RKQEPVRKSTIEKLDKDEIDLVDLVNTFTSD 62

QY 64 YGAELTANVLRDMGLQEMAGQIQAAATHQ-GSGAAPAGIQAPPSAAK 110
Db AVSVTVDIRGKCNVAEELLENTGOG-----GVSQPEPPVPEP 102

RESULT 13
P74033 PRELIMINARY; PRT; 439 AA.
AC P74033;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hypothetical protein slr0806.
GN SLR0806.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
DR EMBL; D90911; BAA18106.1; -.
DR InterPro; IPR006094; Oxid_FAD_bind.
DR Pfam; PF01565; FAD binding 4; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 439 AA; 47989 MW; 6C2E6F56365FA653 CRC64;

Query Match 8.7%; Score 86.5; DB 16; Length 439;
Best Local Similarity 24.0%; Pred. No. 17;
Matches 42; Conservative 22; Mismatches 60; Indels 51; Gaps 8;

QY 31 PLREGYGR---IPRGALLSMDALDITDKLVSYLETYGAELTANVLRDMGLQEMAGQLOA 87
Db PVHGYGTNGIITEITLPLTPALPWREIVSF-----TNLSAIAFAQN 237

QY 88 ATHQ-GSGAAPAGIQAPP-----QSNAPGLHFIDHRAALIAARTNVEM-----LL 133
Db LAHQDGIYSKEISIQADPIQVFSLSKSYQGAHW-----MVIVSELDWLAFTQLA 290

QY 134 DALYGVKVLTDQYQAVRAEPTNPSKRVKLFSTPAWNTCKDLLQALRESOSYL 188
Db KASKGEIIFEQ-----DQSPGKINLIEF-----NWNHTLLARAVDPSLYL 334

RESULT 14
Q8NQ6 PRELIMINARY; PRT; 828 AA.
AC Q8NQ6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta subunit (EC 6.1.1.20).
GN CGL1390.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;

```

```
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT Complete genomic sequence of Corynebacterium glutamicum ATCC 13032." ;
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF005276; BAB98783.1; -.
DR InterPro; IPR005146; B5.
DR InterPro; IPR005147; B5.
DR DR InterPro; IPR005121; Fdx-AntICB.
DR InterPro; IPR004532; PheT_bact..
DR InterPro; IPR002547; tRNA_bind..
DR Pfam; PF03483; B3_4; 1.
DR Pfam; PF03484; B5; 1.
DR Pfam; PF03147; FDX-ACB; 1.
DR Pfam; PF01588; tRNA_bind; 1.
DR TIGRFAMs; TIGR00472_pheT_bact; 2.
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ SEQUENCE      828 AA;  8861 MW;   E3F1B39CE5F29FE CRC64;

    Best Match              8.7%; Score 86.5; DB 16; Length 828;
    Query Local Similarity 24.5%; Pred. No. 41;
    Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7


QY 29 SVPLREGYGRIIPRGALLSMDALDITDKLVSPYL---ETYGAELTANVLRDMGLQEAWGOL 85
Db :: ||| |::||| :|:|:|:|:~::~:||:~::||:~::||:~::||:~::||:~::||:~::
QY 107 AISARETYGRNSAGMICSAESELGLADKQSGIITLDPSYG-EPGEDARQAQLGIEDTVFDV 165
Db ::|||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::

QY 86 QOATHQGGSAPAGIAQPQSAAAPGLHFIDHPPAALIARTNVWLLDALLYGVLT---142
Db :||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::

QY 166 NVTPDPKGVALSARGUTRELASAF--SLTFSTD---PAIEPAVAGIEVKVPVGESLINVEL 220
Db :||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::

QY 143 DSOYOQVAERAEPTNPCKMKLFSTPMWN---WTCKDLLLQALR 182
Db :||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::

QY 221 RRETKAIR-----FGLRKVSGIDPAAESPFWMQRELMLSGQR 257
Db :||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::||:~::

RESULT 15
Q8TSR7 PRELIMINARY;          PRT;       580 AA.

AC AC Q8TSR7 ID Q8TSR7
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE DNA ligase (ATP).
GE LG OR NA0728.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarinales;
OC Methanosarcinaceae; Methanosarcina.
NCBI_TaxID=2214;

[1_]
SEQUENCE FROM N.A.
STRAIN=CZA / ATCC 35395 / DSM 2834;
MEDLINE=21929760; PubMed=11932238;
Galagan J.E., Nusbaum C., Roy A., Smirnov S., Atnoor D., Brown A.,
Fitzhugh W.B., Calvo S., Engels R., Emrizzzi M.G., Macdonald P.,
Allen N., Naylor J., Stange-Thomann K., Dearellano K. Johnson R.,
Linton L., McEwan P., McKernan K., Talamas J.J., Tirrell A., Ye W.W.,
Zimmer A., Barber R.D., Cann I., Graham D.E., Grahané D.A., Guss A.M.,
Heiderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
Springer T.A., Unmaym L.A., White O., White R.H., de Macario E.C.,
Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
Metcalf W.W., Birren B.;
"The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity.";
Nat Genome Res. 12:532-542(2002).

DR EMBL; AE010734; BAM04168.1; -.
DR InterPro; IPR000977; DNA_ligase.
DR DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF01068; DNA_ligase; 1.
DR Pfam; PF04679; DNA_ligase A C; 1.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:11:31 ; Search time 3.84236 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARDALDALENLTAEEL.....LLQALRESQSYLVEDLERS 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pdp:*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pdp:*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pdp:*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pdp:*
- 5: /cgn2_6/prodata/1/iaa/PCUS COMB.pdp:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	4	US-09-340-620A-49
2	684	69.1	193	4	US-09-340-620A-61
3	378	38.2	71	4	US-09-340-620A-58
4	264	26.7	70	4	US-09-340-620A-57
5	264	26.7	70	4	US-09-340-620A-66
6	151.5	15.3	109	4	US-09-340-620A-71
7	94.5	9.5	1233	4	US-09-328-352-7874
8	86.5	8.7	372	1	US-07-813-584A-3
9	86.5	8.7	372	1	US-08-330-515-3
10	84.5	8.5	284	4	US-09-069-023-5
11	84.5	8.5	478	4	US-09-069-023-4
12	84.5	8.5	530	4	US-09-069-023-3
13	84.5	8.5	531	4	US-09-069-023-1
14	84.5	8.5	540	3	US-09-019-942-1
15	84.5	8.5	540	4	US-09-099-041A-2
16	84.5	8.5	540	4	US-09-069-023-27
17	84.5	8.5	540	4	US-09-245-281-2
18	84.5	8.5	540	4	US-09-470-271-1
19	84.5	8.5	540	4	US-09-207-359B-2
20	84.5	8.5	540	4	US-09-340-620A-2
21	84.5	8.5	540	4	US-09-345-473E-28
22	78.5	7.9	424	4	US-09-107-532A-6238
23	77.5	7.8	2436	3	US-08-444-818-75
24	76.5	7.7	164	4	US-09-245-281-41
25	76.5	7.7	164	4	US-09-207-359B-41
26	76.5	7.7	164	4	US-09-340-620A-41
27	76.5	7.7	249	4	US-09-245-281-39

ALIGNMENTS

RESULT 1

US-09-340-620A-49

; Sequence 49, Application US/09340620A

; Patent No. 6482933

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF

; FILE REFERENCE: 07334-124001

; CURRENT APPLICATION NUMBER: US/09/340,620A

; CURRENT FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: US 09/245,281

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: US 09/207,359

; PRIOR FILING DATE: 1998-12-08

; PRIOR APPLICATION NUMBER: US 09/099,041

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: US 09/019,942

; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 49

; LENGTH: 195

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-340-620A-49

Query Match 100.0%; Score 990; DB 4; Length 195;

Best Local Similarity 100.0%; Pred. No. 1.8e-112;

Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDALDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMALDLDTKLVSPY 60

Db 1 MGRARDALDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMALDLDTKLVSPY 60

QY 61 LETYCAELTANVLRDMGLQEMAGOLQAATHOGSGGAAPAGIQAPPQSAKPGHLHFIDQHRA 120

Db 61 LETYCAELTANVLRDMGLQEMAGOLQAATHOGSGGAAPAGIQAPPQSAKPGHLHFIDQHRA 120

QY 121 ALIARVTNVEMLLDALYGLKVLTDQYQAVRAEPTNPSKMKLFSTPAWNTWKDILLQA 180

Db 121 ALIARVTNVEMLLDALYGLKVLTDQYQAVRAEPTNPSKMKLFSTPAWNTWKDILLQA 180

QY 181 LRESQSYLVEDLERS 195

Db 181 LRESQSYLVEDLERS 195

RESULT 2

US-09-340-620A-61

; Sequence 61, Application US/09340620A

```
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-61

Query Match
Best Local Similarity 69.1%; Score 684; DB 4; Length 193;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSPY 60
DB 1 MGRARDAILDENLTAEELKFKFKMLLTVOLREGYGRIPRGALLQMDAIDLTKLVSY 60
QY 61 LETYGAELTAVLRDMLGLOSMAGLOQAATHOGSGAAPAGIQAPPOQAAKPGHLHFDIOHRA 120
DB 61 LESYGLELTAVLRDMLGLOLQELAEQLQ-TTYESGAVAAAASVPAQSTARTG-HFVDQHRQ 118
QY 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTFPANWTKDQLLQA 180
DB 119 ALIARTVTEVDGLDALHGSVLTGEGYQAVRAETTSQDKMKLFSTFPVSNWLNCKDQLLQA 178
QY 181 LRESQSYVEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193

RESULT 3
US-09-340-620A-58
; Sequence 58, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-58

Query Match
Best Local Similarity 38.2%; Score 378; DB 4; Length 71;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 111 GLHFIDQHRAALIAARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTFPANW 170
DB 1 GLHFIDQHRAALIAARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTFPANW 60
QY 171 WTKDQLLQAL 181
DB 61 WTKDQLLQAL 71
```

```
RESULT 4
US-09-340-620A-57
; Sequence 57, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-57
```

```
Query Match
Best Local Similarity 26.7%; Score 264; DB 4; Length 70;
Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 113 HFIDQHRAALIAARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTFPANW 172
DB 2 HFVDQHRAALIAARTVTEVDGLDALHGSVLTGEGYQAVRAETTSQDKMKLFSTFPVSNW 61
QY 173 CKDQLLQAL 181
DB 62 CKDQLLQAL 70
```

```
RESULT 5
US-09-340-620A-66
; Sequence 66, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1999-02-05
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
```

```

US-09-340-620A-66
Query Match      26.7%; Score 264; DB 4; Length 70;
Best Local Similarity 73.9%; Pred. No. 1.2e-24;
Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 113 HFIDQRAALIAITVNEVLLDALYGVKLTDEQYQAVRAEPTNPSKMKRLFSFTFAWNT 172
DB 2 HFVDHQRAIALIARVTEVDGVLGDALHGSVLTQGYQAVRAETTSQDKMKRLFSFVSNLIT 61
QY 173 CKDLLLQAL 181
DB 62 CKDSLQAL 70

RESULT 6
US-09-340-620A-71
Sequence 71, Application US/09340620A
Patent No. 6482933
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-124001
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US/09/340,620A
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 71
LENGTH: 109
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(109)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-340-620A-71

Query Match      15.3%; Score 151.5; DB 4; Length 109;
Best Local Similarity 46.7%; Pred. No. 1.2e-10;
Matches 42; Conservative 13; Mismatches 28; Indels 7; Gaps 4;

QY 111 GLHFIDQRAALIAITVNT--VEWLLDALYGVK-VLTDEQYQAVRAEPTNPSKMKRLFSFTFP 167
DB 7 GSEIDQHXALLARVTEPDPSLLDALLSRDLISEDYEAETTXLSKVRKLLILVQ 66
QY 168 A-WNWTCKDL---LLQALRESOSYLVEDLE 193
DB 67 SXGEETCKXFLKCLLQALKASAAVGLDPE 96

RESULT 7
US-09-328-352-7874
Sequence 7874, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 9252
SEQ ID NO 7874
LENGTH: 1233

```

```

; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7874

Query Match      9.5%; Score 94.5; DB 4; Length 1233;
Best Local Similarity 26.5%; Pred. No. 0.047; 76; Indels 45; Gaps 10;
Matches 53; Conservative 26; Mismatches 26; Indels 45; Gaps 10;

QY 8 ILDALENITABELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFYLETYGAE 67
DB 125 VIDQLDELFGVTLDSFSQKLLREFAFES-GKIER-AQITDDA-----KTYSRQ 170
QY 68 LTANVLRD-----MGLQEMAGLOQAA-----THQSGAAPAGIQAPQSAKPG 112
DB 171 LTHDVLREWIQSPQQTVIDALYLAGEKLSVDSFVKLVEDSLNFSAHFKLP---EKPTI 226
QY 113 HFIDQRAALIAITVNEV--WLLDALYGVKLTDEQYQAVRAEPTNPSKMKRLFSFTFAWNT 171
DB 227 QF--EQLAQLKQLATEIDISLEPY--LLDGEHYKHVNGTIFRNGAFNKLF----- 275
QY 172 TKCOLLLQALRESOSYLVED 191
DB 276 ECLPQLQLKQSDSILVFD 295

RESULT 8
US-07-813-584A-3
Sequence 3, Application US/07813584A
Patent No. 5352588
GENERAL INFORMATION:
APPLICANT: Fischetti, Vincent A.
APPLICANT: Bessen, Debra E.
TITLE OF INVENTION: No. 5352588e1 Immunoglobulin A Binding Protein
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kittie Murray
STREET: 98 Cutter Mill Road
CITY: Great Neck
STATE: NY
COUNTRY: USA
ZIP: 11021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/813,584A
APPLICATION NUMBER: US/07/813,584A
FILING DATE: 19911224
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Kittie
REGISTRATION NUMBER: 30,246
REFERENCE/DOCKET NUMBER: RU-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-482-1990
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-813-584A-3

Query Match      8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 0.068;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

QY 12 LENLTABELKKFKLLSVPLREGYGRIPRGALLSMDAL-----DLTDKLVSFYLETYGA 66
DB 78 LEKINAEENKKNLEAKINKELNENYKLDG-----IDALEKEKEDLKTTLAKTTENEIS 133
QY 67 ELTANVL-RDMGLQEMAGLOQAATHQSGAAPAGIQAPPQ--SAAKPGL-HFIDQRAAL 122

```

Db 134 EASRGLSRDLEASRTAKKELEAKHQLEAKNKLTEGNQVSEASRKGSLNDLEASRAA- 192
QY 123 IARVTNVEWLLDALYGVLTDEQ-----YQAVRAEPTNFSKMRKLPSPFPANNWTKDLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHQLEADYQVSETSRKGLS-----RD-- 234
QY 178 LQALRESQSVLVDLERS 195
Db 235 LEASREANKVTSelta 252

RESULT 9
US-08-330-515-3
; Sequence 3, Application US/08330515
; Patent No. 5556944
; GENERAL INFORMATION:
; APPLICANT: Fischetti, Vincent A.
; APPLICANT: Bessen, Debra E.
; TITLE OF INVENTION: No. 5556944el Immunoglobulin A Binding Protein
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,515
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,584
; FILING DATE: 24-DEC-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: RU-100.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-330-515-3

Query Match 8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 0.068;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;
QY 12 LENLTAEELKFKLLSVLPREGYGRIPRGALLSMDAL-----DLTDKLVSYFLEYGA 66
Db 78 LEKINAEENKVKLEAINELNENYVKLODG-----IDALEKEKEDLKTTLAKTKENEIS 133
QY 67 ELTANYLV-RDMGLQEWAGLOQAATHOGSGAAPAGIQAPPQ--SAAKPGL-HFIDQHRAAL 122
Db 134 EASRGLSRDLEASRTAKKELEAKHQLEAKNKLTEGNQVSEASRKGSLNDLEASRAA- 192
QY 123 IARVTNVEWLLDALYGVLTDEQ-----YQAVRAEPTNFSKMRKLPSPFPANNWTKDLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHQLEADYQVSETSRKGLS-----RD-- 234
QY 178 LQALRESQSVLVDLERS 195

Db 235 LEASREANKVTSelta 252
RESULT 10
US-09-069-023-5
; Sequence 5, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-5
Query Match 8.5%; Score 84.5; DB 4; Length 284;
Best Local Similarity 20.1%; Pred. No. 0.079;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;
QY 6 DAILEDALNLTAE-----LKKFKLLSVLPREGYGRIPRGALLSMDALDITDKLVSYFL 61
Db 35 EPVLTPTPEITFLEAVIQLKTKTKLQSVS-----SAIHLCDK----- 70
QY 62 ETYGAELTANVLDMGLQ-ENAGLOQAATHOGSGAAPAGIQAP----- 103
Db 71 --KKMELSLNIPVNHGPOECSGSQLHENSGLSPAPQDNDFLSRAQDCYPMK 128
QY 104 -----POSAA-----KPGDL--HFIDQHR 119
Db 129 LHHCPGNHSDSTISGSORAAFCDBKHTTCCSALINPLSTAGNSERLOPGIAQQWIOSKR 188
QY 120 AALITARVTN--VEWLLDALYGVLTDEQYQAVRAEPTNFSKMRKLPSPFPANNWTKDL 176
Db 189 EDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLDITDIQGEFAKV 248
QY 177 LLOALRESQ 185
Db 249 IVQKLKDNK 257

RESULT 11
US-09-069-023-4
; Sequence 4, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-4
Query Match 8.5%; Score 84.5; DB 4; Length 478;
Best Local Similarity 20.1%; Pred. No. 0.18;

```

; ORGANISM: Homo sapiens
US-09-069-023-3

Query Match      8.5%; Score 84.5; DB 4; Length 530;
Best Local Similarity 20.1%; Pred. No. 0.21;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9

Qy      6 DAILDALENLTAE-----LKKFKLLSVPLREGYGRIPRGALLISMDALDITDKLVSYFL 61
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      281 EPVLRTFEEITFLEAVIQLKTKLQSVS-----SAIHLCDK----- 316
      : : : : : : : : : : : : : : : : : : : : : : : :

Qy      62 ETYGAELTANVLRDWMGLQ-EMAGOLQAATHOGSGAAPAGIOAP----- 103
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      317 --KQWELSLNIPVNHGPOEESCGSQLHENGSGSPETSRLPAPQDNDFLSRKAQDCYPMK 374
      : : : : : : : : : : : : : : : : : : : : : : : :

Qy      104 -----POSAA-----KPGL--HFIDQHR 119
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      375 LHHCPGNHMSWDSTIGSQRAAFCDHKHTTFCSSAIINPLSTAGNSERLQPGIAQQWIOQSR 434
      : : : : : : : : : : : : : : : : : : : : : : : :

Qy      120 AALIRVTN--VEWLLDALYCK-VLTDROYQAVRAEPTNPISKVKLFSTPAWNWTKDL 176
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      435 EDIVNQWTEACNOSLDALLSRDLIMKEDYELVSTKPTRTSKVROLLDTTDIQGEPAKV 494
      : : : : : : : : : : : : : : : : : : : : : : : :

Qy      177 LQALRESQ 185
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      495 IVQKLKDNK 503
      : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-09-069-023-1
; Sequence 1, Application US/09069023A
; Patent No. 6348573

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 ; Search time 33.9809 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTABEL.....LLQLALRSQSYLVEDLRS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues
Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	9	US-09-728-721-49
2	990	100.0	195	10	US-09-996-617-8
3	990	100.0	195	10	US-09-841-879B-5
4	990	100.0	195	12	US-10-131-410-84
5	990	100.0	195	12	US-10-240-145-55
6	990	100.0	195	15	US-10-295-981-49
7	990	100.0	205	15	US-10-106-698-5421
8	759	76.7	190	9	US-09-925-301-1120
9	684	69.1	193	9	US-09-728-721-61
10	684	69.1	193	10	US-09-841-879B-2
11	684	69.1	193	15	US-10-295-981-61
12	564.5	57.0	158	15	US-10-106-698-5422
13	509	51.4	136	9	US-09-925-299-1421
14	509	51.4	136	11	US-09-925-299-1421
15	469	47.4	90	10	US-09-931-071-7

16	445	44.9	85	10	US-09-841-879B-8	Sequence 8, Appli
17	435	43.9	89	14	US-10-127-516-16	Sequence 16, Appli
18	435	43.9	89	14	US-10-027-629-16	Sequence 16, Appli
19	378	38.2	71	9	US-09-728-721-58	Sequence 58, Appli
20	378	38.2	71	15	US-10-295-981-58	Sequence 58, Appli
21	377	38.1	77	14	US-10-127-516-8	Sequence 8, Appli
22	377	38.1	77	14	US-10-027-629-8	Sequence 8, Appli
23	370	37.4	77	12	US-10-132-967-8	Sequence 9, Appli
24	368	37.2	76	11	US-09-965-621-9	Sequence 9, Appli
25	368	37.2	76	12	US-10-407-866-9	Sequence 9, Appli
26	309	31.2	85	10	US-09-841-879B-15	Sequence 15, Appli
27	301	30.4	84	10	US-09-841-879B-7	Sequence 7, Appli
28	267	27.0	89	11	US-09-965-621-28	Sequence 28, Appli
29	267	27.0	89	12	US-10-407-866-28	Sequence 28, Appli
30	267	27.0	89	14	US-10-127-516-17	Sequence 17, Appli
31	267	27.0	89	14	US-10-027-629-17	Sequence 17, Appli
32	264	26.7	70	9	US-09-728-721-57	Sequence 57, Appli
33	264	26.7	70	9	US-09-728-721-66	Sequence 66, Appli
34	264	26.7	70	15	US-10-295-981-57	Sequence 57, Appli
35	264	26.7	70	15	US-10-295-981-66	Sequence 66, Appli
36	250.5	25.3	1399	10	US-09-388-221-4	Sequence 4, Appli
37	250.5	25.3	1429	10	US-09-996-617-2	Sequence 2, Appli
38	250.5	25.3	1429	10	US-09-931-071-2	Sequence 2, Appli
39	250.5	25.3	1429	12	US-10-028-374-15	Sequence 15, Appli
40	250.5	25.3	1429	12	US-10-183-770-15	Sequence 15, Appli
41	250.5	25.3	1429	15	US-10-028-392-11	Sequence 11, Appli
42	250	25.3	1443	10	US-09-388-221-6	Sequence 6, Appli
43	250	25.3	1473	10	US-09-388-221-2	Sequence 2, Appli
44	249.5	25.2	1429	12	US-10-028-374-3	Sequence 3, Appli
45	249.5	25.2	1429	12	US-10-183-770-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-728-721-49
; Sequence 49, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-721-49

Query Match 100.0%; Score 990; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGRARDAILDALENLTABELKFKLLSVLPUREGYGRIPRGALLSMDALDLDKLVSY	60
Db	1	MGRARDAILDALENLTABELKFKLLSVLPUREGYGRIPRGALLSMDALDLDKLVSY	60
Qy	61	LETYGAEILTANVLRDMGLQEMAGLOAATHQSGAPAGIQAPPOSAAKPGHFIIDQHRA	120
Db	61	LETYGAEILTANVLRDMGLQEMAGLOAATHQSGAPAGIQAPPOSAAKPGHFIIDQHRA	120

[illegible]

```
RESULT 5
US-10-240-145-55
; Sequence 55, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/340,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 55
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145-55

Query Match      100.0%; Score 990; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
QY 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
DB 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 6
US-10-295-981-49
; Sequence 49, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
QY 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
DB 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 7
US-10-106-698-5421
; Sequence 5421, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5421
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5421

Query Match      100.0%; Score 990; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 70
QY 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
DB 71 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 130
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
DB 131 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 190
QY 181 LRESQSYLVEDLERS 195
DB 191 LRESQSYLVEDLERS 205
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-981-49

Query Match      100.0%; Score 990; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
QY 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
DB 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195
```

```
Query Match      100.0%; Score 990; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 60
DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSYF 70
QY 61 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 120
DB 71 LETYGAELTANVLDRMDGLQEMAGOLQAATHQSGGAAPAGIQAPPSAAKPGHLHFIDQHRA 130
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 180
DB 131 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPPAWNWTCKDILLQA 190
QY 181 LRESQSYLVEDLERS 195
DB 191 LRESQSYLVEDLERS 205
```

RESULT 8

US-09-925-301-1120
; Sequence 1120, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1120

Query Match 76.7%; Score 759; DB 9; Length 190;
Best Local Similarity 95.6%; Pred. No. 2.1e-73;
Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60

DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 70

QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPQSAAPGLHFDQHRA 120

DB 71 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPQSAAPGLHFDQHRA 130

QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMR 160

DB 131 ALIARVTNVWLLDALYGVLTDEQYQAVR--PSPPTQAR 168

RESULT 9

US-09-728-721-61
; Sequence 61, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-728-721-61

Query Match 69.1%; Score 684; DB 9; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60

DB 1 MGRARDAILDALENLSGDELKKFKMLLTQVLRREGYGRIPRGALLQMDAIDLTDKLVSY 60

QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPQSAAPGLHFDQHRA 120
DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKESGAVAAAAASVPAQSTARTG-HFVDQHRQ 118
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDILLQA 180
DB 119 ALIARVTEVDGVLDAHGHVLTGEGYQAVRAETTSQDKMRKLFSPVPSMNLTKDLSLLQA 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193

RESULT 10

US-09-841-879B-2
; Sequence 2, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-130001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-841-879B-2

Query Match 69.1%; Score 684; DB 10; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60

DB 1 MGRARDAILDALENLSGDELKKFKMLLTQVLRREGYGRIPRGALLQMDAIDLTDKLVSY 60

QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPQSAAPGLHFDQHRA 120

DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKESGAVAAAAASVPAQSTARTG-HFVDQHRQ 118

QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDILLQA 180

DB 119 ALIARVTEVDGVLDAHGHVLTGEGYQAVRAETTSQDKMRKLFSPVPSMNLTKDLSLLQA 178

QY 181 LRESQSYLVEDLERS 195

DB 179 LKEIHPYLVMDLEQS 193

RESULT 11

US-10-295-981-61
; Sequence 61, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041

```

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-295-981-61

Query Match
  69.1%; Score 684; DB 15; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60

QY 61 LETYGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPAGIQAPPSAAKPGHLFIDQHRA 120
DB 61 LESYGLELTWVLRDMGLOELAEQLQ-TTKESGAVAAAAAASVPAQSTARTG-HFVDQHRQ 118

QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSPFTPAWNTCKDLLLQA 180
DB 119 ALIARVTEVDGVLDAHGSVLTGQYQAVRAETTSQDKMRKLFSPVPSWNLTKDLSLQA 178

QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPYLVNDLEQS 193

RESULT 12
US-10-106-698-5422
; Sequence 5422, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5422
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5422

Query Match
  57.0%; Score 564.5; DB 15; Length 158;
Best Local Similarity 75.8%; Pred. No. 1.4e-52;
Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 ARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFYLET 63
DB 13 ARDAILDALENLTAEELKKE-----KLVSFYLET 41

QY 64 YGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPAGIQAPPSAAKPGHLFIDQHRAALI 123
DB 42 YGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPAGIQAPPSAAKPGHLFIDQHRAALI 101

QY 124 ARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMR 160
DB 102 ARVTNVWLLDALYGVLTDEQYQAVR--PSPPTQAR 136
```

```

RESULT 13
US-09-925-299-1421
; Sequence 1421, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1421
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match
  51.4%; Score 509; DB 9; Length 136;
Best Local Similarity 83.7%; Pred. No. 1e-46;
Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 70

QY 61 LETYGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPAG-----IQAPPSAAKPGHLFID 116
DB 71 LETYGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPLGSRPLLSRQPSOACT----LID 126

QY 117 QHRAALIR 125
DB 127 QHRAALSRR 135

RESULT 14
US-09-925-299-1421
; Sequence 1421, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1421
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match
  51.4%; Score 509; DB 11; Length 136;
Best Local Similarity 83.7%; Pred. No. 1e-46;
Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 70

QY 61 LETYGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPAG-----IQAPPSAAKPGHLFID 116
DB 71 LETYGAELTANVLRDMGLOEMAGLOAQATHQSGGAAPLGSRPLLSRQPSOACT----LID 126
```

QY 117 QHRAALIAIAR 125
| | | | |
Db 127 QHRAALSRR 135

RESULT 15

US-09-931-071-7
; Sequence 7, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-071-7

Query Match 47.4%; Score 469; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.2e-42;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 SAAKPGIHFIDQHRAALIAIARVTNVVWLLDALYGKVLTDQYQAVRAEPTNPSSKQKLFSP 165
| | | | |
Db 1 SAAKPGIHFIDQHRAALIAIARVTNVVWLLDALYGKVLTDQYQAVRAEPTNPSSKQKLFSP 60
| | | | |
QY 166 TPANNWTKDILLQALRESQSILVEDLERS 195
| | | | |
Db 61 TPANNWTKDILLQALRESQSILVEDLERS 90
| | | | |

Search completed: January 29, 2004, 13:54:38
Job time : 34.9809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 19587.5 Seconds
(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gcccagggctggagaggt.....aggaataagaagttacctac 5444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sta:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sta:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pin:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sv:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5444	100.0	5444	9	AB023143	AB023143 Homo sapi
2	5400.8	99.2	5404	9	AF310105	AF310105 Homo sapi
3	4882.4	89.7	5100	6	AX459865	AX459865 Sequence
4	4882.4	89.7	5100	6	AF298548	AF298548 Homo sapi
5	4765.6	87.5	5122	6	AX207661	AX207661 Sequence
6	4290	78.8	4290	9	AF229059	AF229059 Homo sapi
7	4281.8	78.7	5306	9	BC051787	BC051787 Homo sapi
8	4148	76.2	4422	6	AX089761	AX089761 Sequence
9	4148	76.2	4422	9	AF229060	AF229060 Homo sapi
10	4100	75.3	4200	9	AF229061	AF229061 Homo sapi
11	4078	74.9	4194	6	AX089763	AX089763 Sequence
12	3958	72.7	4332	9	AF229062	AF229062 Homo sapi
13	3952	72.6	4329	6	AX089765	AX089765 Sequence
14	3807.8	69.9	4123	9	HSM800983	AL117470 Homo sapi
15	3400.4	62.5	4556	6	AX089769	AX089769 Sequence
16	3210.4	59.0	4466	6	AX089771	AX089771 Sequence
17	3094.8	56.8	3419	9	AK026393	AK026393 Homo sapi
18	1710.8	31.4	152812	2	AC090536	AC090536 Homo sapi
19	1710.8	31.4	176730	9	AC055839	AC055839 Homo sapi
20	1566.4	28.8	2373	9	AK026398	AK026398 Homo sapi
21	991.8	18.2	111862	2	HSAC001237	AC001237 Homo sapi
22	954	17.5	111862	2	HSAC001237	AC001237 Homo sapi
23	742.8	13.6	174564	2	AL663042	AL663042 Mouse DNA
24	738	13.6	174564	2	AC121704	AC121704 Rattus no
25	738	13.6	250275	2	AC095695	AC095695 Rattus no
26	728	13.4	221540	2	AC026912	AC026912 Mus muscu
27	722	13.3	239196	2	AC127967	AC127967 Rattus no
28	701.6	12.9	96115	10	AL662908	AL662908 Mouse DNA
29	695.2	12.8	180274	2	AC027185	AC027185 Mus muscu
30	603	11.1	254690	2	AC095179	AC095179 Rattus no
31	399.8	7.3	3827	9	AY095146	AY095146 Homo sapi
32	399.8	7.3	3827	9	AY154467	AY154467 Homo sapi
33	399.8	7.3	69124	2	AC069567	AC069567 Mus muscu
34	398.2	7.3	3466	9	AK095460	AK095460 Homo sapi
35	398.2	7.3	3507	9	BC028069	BC028069 Homo sapi
36	398	7.3	3108	6	AX417214	AX417214 Sequence
37	398	7.3	3300	6	AX459869	AX459869 Sequence
38	398	7.3	3563	9	AY116205	AY116205 Homo sapi
39	398	7.3	3731	9	AY116204	AY116204 Homo sapi
40	397.4	7.3	3395	9	AY116206	AY116206 Homo sapi
41	395.8	7.3	3221	9	AY116207	AY116207 Homo sapi
42	392.2	7.2	108555	2	AC079221	AC079221 Mus muscu
43	376	6.9	3828	9	AF427617	AF427617 Homo sapi
44	374.4	6.9	3237	9	AY092033	AY092033 Homo sapi
45	373	6.9	4170	9	AF410477	AF410477 Homo sapi

ALIGNMENTS

RESULT 1
AB023143
LOCUS
DEFINITION Homo sapiens mRNA for KIAA0926 protein, complete cds.
ACCESSION AB023143
VERSION AB023143.1 GI:4589483
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M.,
Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
TITLE

5444 bp mRNA linear PRI 16-JUN-1999

XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
DNA Res. 6 (1), 63-70 (1999)
99246063
10231032
2 (bases 1 to 5444)
Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (04-FEB-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan [E-mail: cdnainfo@kazusa.or.jp. Tel: +81-438-52-3913,
Fax: +81-438-52-3914]

BASE COUNT	1303	a	1525	c	1474	g	1142	t	
ORIGIN									
Query Match									100.0%; Score 5444; DB 9; Length 5444;
Best Local Similarity									100.0%; Pred. NO. 0;
Matches 5444; Conservative									0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	GCC	CCAGGGG	CTCGAGAG	GTCTGAAGAA	AACTCGGAG	CCAGCAG	CCCGGGCTCCACT	CTCT 60
Db	1	GCC	CCAGGGC	CTGGAGNG	GTCTGAGAA	AACTCGGAG	CCAGCAG	CCCGGGCTCCACT	CTCT 60
Qy	61	GG	TTCTGA	AAAGCCCA	TTCCTG	CTCGGGCT	CTCCAC	CCCCACCTCTT	TCAGCCTT 120
Db	61	GG	TTCTGA	AAAGCCCA	TTCCTG	CTCGGGCT	CTCCAC	CCCCACCTCTT	TCAGCCTT 120
Qy	121	GCAG	CTCAAGGG	TGTATCT	TCAGGAG	TCCAGGAC	CCCGAGGAG	GGAAGAAATCT	TCAGGAACA 180
Db	121	GCAG	CTCAAGGG	TGTATCT	TCAGGAG	TCCAGGAC	CCCGAGGAG	GGAAGAAATCT	TCAGGAACA 180

Qy	181	CAGAA	CAGT	GAG	GGT	TGCC	CAC	ACCC	CAAT	CTCC	GTCA	CCACA	TCTCC	GTCA	CCACA	TCTCC	GTCA	CCACT	CAC	240	
Db	181	CAGAA	CAGT	GAG	GGT	TGCC	CAC	ACCC	CAAT	CTCC	GTCA	CCACA	TCTCC	GTCA	CCACA	TCTCC	GTCA	CCACT	CAC	240	
Qy	241	CTCT	CCCT	GTG	CTG	CCCT	GTG	ACCC	CAAT	CTCC	GTG	ACCC	CAAT	CTCC	GTG	ACCC	CAAT	CTCC	GTG	300	
Db	241	CTCT	CCCT	GTG	CTG	CCCT	GTG	ACCC	CAAT	CTCC	GTG	ACCC	CAAT	CTCC	GTG	ACCC	CAAT	CTCC	GTG	300	
Qy	301	GTCT	TTC	GAG	CCCT	CTG	GGCT	CCT	CCCT	CCCT	GGCT	TTT	TCT	TAC	CACT	CCCT	CCCT	CTAT		360	
Db	301	GTCT	TTC	GAG	CCCT	CTG	GGCT	CCT	CCCT	CCCT	GGCT	TTT	TCT	TAC	CACT	CCCT	CCCT	CTAT		360	
Qy	361	CGCG	CTCT	ATCT	GTAG	TG	CCCT	GTG	GAAT	TATA	AAACT	CTGG	TTT	CGAA	TCT	CGAA	TAAGA			420	
Db	361	CGCG	CTCT	ATCT	GTAG	TG	CCCT	GTG	GAAT	TATA	AAACT	CTGG	TTT	CGAA	TCT	CGAA	TAAGA			420	
Qy	421	GACG	TTAAG	CCCAAG	GCAC	AGCA	CTG	TTCT	TG	CCCT	GTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG		480	
Db	421	GACG	TTAAG	CCCAAG	GCAC	AGCA	CTG	TTCT	TG	CCCT	GTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG		480	
Qy	481	CTGG	GAAC	TCC	CC	CAG	ACA	CCCT	CTT	AACT	CC	GG	GAC	GAG	AT	TGG	CTGG	CG	AG	CTG	540
Db	481	CTGG	GAAC	TCC	CC	CAG	ACA	CCCT	CTT	AACT	CC	GG	GAC	GAG	AT	TGG	CTGG	CG	AG	CTG	540
Qy	541	GGCG	CTG	CCCT	GTG	TACT	TGG	AGTT	CTCT	GA	AG	AG	GG	AG	GCT	GA	AG	AGTT		600	
Db	541	GGCG	CTG	CCCT	GTG	TACT	TGG	AGTT	CTCT	GA	AG	AG	GG	AG	GCT	GA	AG	AGTT		600	
Qy	601	CTGCT	CG	CAAT	TAAG	CG	CACT	CTC	G	AG	GCT	CTT	CG	GT	G	AG	CA	CA	CGCT		660
Db	601	CTGCT	CG	CAAT	TAAG	CG	CACT	CTC	G	AG	GCT	CTT	CG	GT	G	AG	CA	CA	CGCT		660
Qy	661	AAG	AC	CAG	TGG	CA	TG	AG	GT	GG	CT	CG	T	AG	T	GG	G	AG	CAG		720
Db	661	AAG	AC	CAG	TGG	CA	TG	AG	GT	GG	CT	CG	T	AG	T	GG	G	AG	CAG		720
Qy	721	TGG	AG	CT	T	AG	CCCT	CA	T	AC	T	G	G	G	A	G	A	G	G		780
Db	721	TGG	AG	CT	T	AG	CCCT	CA	T	AC	T	G	G	G	A	G	A	G	G		780
Qy	781	CAG	GA	AG	GG	CG	AG	CA	CT	CT	CC	CT	CA	T	T	CC	CT	CA	AG		840
Db	781	CAG	GA	AG	GG	CG	AG	CA	CT	CT	CC	CT	CA	T	T	CC	CT	CA	AG		840
Qy	841	TCT	CC	CAG	CCAA	CC	CA	CT	CC	AC	CG	CAG	TG	CT	AA	TG	CC	T	GA		900
Db	841	TCT	CC	CAG	CCAA	CC	CA	CT	CC	AC	CG	CAG	TG	CT	AA	TG	CC	T	GA		900
Qy	901	GG	GT	CA	CC	AG	GG	CT	C	A	G	A	G	GG	TTT	T	G	A	G		960
Db	901	GG	GT	CA	CC	AG	GG	CT	C	A	G	A	G	GG	TTT	T	G	A	G		960
Qy	961	CGT	T	G	A	G	A	A	A	T	CT	T	G	C	T	CA	T	C	T		1020
Db	961	CGT	T	G	A	G	A	A	A	T	CT	T	G	C	T	CA	T	C	T		1020
Qy	1021	GAG	T	CT	CCA	AG	CC	CAG	AG	T	CA	CC	CA	CG	CC	CC	CA	T	CA		1080
Db	1021	GAG																			

QY	1261	AGCTACAGCCCAACCAACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTGTCACAA	1320
DB	1261	AGCTACAGCCCAACCAACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTGTCACAA	1320
QY	1321	TGGCCCTGAAATAAGAGATTTTAAACAAATTAACACAGCTGCTACTCTCAAAAGA	1380
DB	1321	TGGCCCTGAAATAAGAGATTTTAAACAAATTAACACAGCTGCTACTCTCAAAAGA	1380
QY	1381	CCTCACCCAGAGCAATTAATGAGATCAAGAGAACTGTCCTGATTTATGTGGAGAG	1440
DB	1381	CCTCACCCAGAGCAATTAATGAGATCAAGAGAACTGTCCTGATTTATGTGGAGAG	1440
QY	1441	AATCGAGGACATTTAATGAGATCAAGAGAACTGTCCTGATTTATGTGGAGAG	1500
DB	1441	AATCGAGGACATTTAATGAGATCAAGAGAACTTATTGGCCAGGCTGATACCCAAAGA	1500
QY	1501	CCTCGCATAGTCAATCTGAGGGGGCTGCTGGAATTTGGGAATCAACACTGCCAGGAG	1560
DB	1501	CCTCGCATAGTCAATCTGAGGGGGCTGCTGGAATTTGGGAATCAACACTGCCAGGAG	1560
QY	1561	GTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCAGCATGCTTCTAC	1620
DB	1561	GTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCAGCATGCTTCTAC	1620
QY	1621	TTGAGCTGAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTCGCTGAGCTCATCGGAAA	1680
DB	1621	TTGAGCTGAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTCGCTGAGCTCATCGGAAA	1680
QY	1681	GATGGGACAGCACTCCGGCTCCCATTAGACAGATCCTCTTAGGCCAGAGCGGCTGCTC	1740
DB	1681	GATGGGACAGCACTCCGGCTCCCATTAGACAGATCCTCTTAGGCCAGAGCGGCTGCTC	1740
QY	1741	TTGATCTCGATGGTGTAGATGAGCAGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC	1800
DB	1741	TTGATCTCGATGGTGTAGATGAGCAGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC	1800
QY	1801	TGTCTGACTGGAGCAGCAGCAGCCGCGGATGCACTGCTGGGCAAGTTGCTGGGAAA	1860
DB	1801	TGTCTGACTGGAGCAGCAGCAGCCGCGGATGCACTGCTGGGCAAGTTGCTGGGAAA	1860
QY	1861	ACTATACCTCCGAGGATCCTTCTGATGAGCAGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC	1920
DB	1861	ACTATACCTCCGAGGATCCTTCTGATGAGCAGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC	1920
QY	1921	ATTCTCTTTTGGAGCAGGACCTTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGG	1980
DB	1921	ATTCTCTTTTGGAGCAGGACCTTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGG	1980
QY	1981	AAGGAATATTTTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG	2040
DB	1981	AAGGAATATTTTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG	2040
QY	2041	GTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGGCCCTGGGTGTCTGGCTGGCC	2100
DB	2041	GTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGGCCCTGGGTGTCTGGCTGGCC	2100
QY	2101	TGCATTTGCTGATGAGCAGATGAAGCGAAGGAAATCACTACCTGCTTCCAGAGCC	2160
DB	2101	TGCATTTGCTGATGAGCAGATGAAGCGAAGGAAATCACTACCTGCTTCCAGAGCC	2160
QY	2161	ACCACAACTCTGTCTACATTTACCTTTGCCAGGCTCTCAAGCTCAGGCCATTGGGACCC	2220
DB	2161	ACCACAACTCTGTCTACATTTACCTTTGCCAGGCTCTCAAGCTCAGGCCATTGGGACCC	2220
QY	2221	CAGCTCAGAGACCTCTGCTCTGCTGCTGAGGGCATCTGGCAAAAAAGACCTTTTTC	2280
DB	2221	CAGCTCAGAGACCTCTGCTCTGCTGCTGAGGGCATCTGGCAAAAAAGACCTTTTTC	2280
QY	2281	AGTCCAGATGACCTCAGGAGAGATGGTTAGATGGGGCATCATCTCCACCTCTTGAAG	2340
DB	2281	AGTCCAGATGACCTCAGGAGAGATGGTTAGATGGGGCATCATCTCCACCTCTTGAAG	2340
QY	2341	ATGGGTATTCTTCAAGAGCACCCTCTGAGCTACAGCTTCACTCTGTTTC	2400

DB	2341	ATGGGTATTCTTCAAGAGCACCCTCTGAGCTACAGTTCACTTCACTCTGTTTC	2400
QY	2401	CAAGAGTTCTTTTGAGCAATGTCTTATCTTTTGAGGATGAGAAGGGGAGAGTTAAACAT	2460
DB	2401	CAAGAGTTCTTTTGAGCAATGTCTTATCTTTTGAGGATGAGAAGGGGAGAGTTAAACAT	2460
QY	2461	TCTAATTCATCATAGATTTGGAAAAAGACGCTAGAAAGCATATGGAATACATGGCCTGTTT	2520
DB	2461	TCTAATTCATCATAGATTTGGAAAAAGACGCTAGAAAGCATATGGAATACATGGCCTGTTT	2520
QY	2521	GGGGCATCAACACACGTTTCTTATTTGGCCCTGTTTAACTGATGAGGGGAGAGAGATG	2580
DB	2521	GGGGCATCAACACACGTTTCTTATTTGGCCCTGTTTAACTGATGAGGGGAGAGAGATG	2580
QY	2581	GAGAACTCTTTTCACTGCGCGCTGTCTCAGGGGAGAACTTGATGACGTGGGTCCCGTCC	2640
DB	2581	GAGAACTCTTTTCACTGCGCGCTGTCTCAGGGGAGAACTTGATGACGTGGGTCCCGTCC	2640
QY	2641	CTGCAGCTGCTGCTGAGCCACACTCTCTGAGTCCCTCCACTGTTGTACGAGACTCGG	2700
DB	2641	CTGCAGCTGCTGCTGAGCCACACTCTCTGAGTCCCTCCACTGTTGTACGAGACTCGG	2700
QY	2701	AACAAAACTTCTGACACAAAGTATGCCCCATTTTGAAGAAATGGGCATGTTGTAGAA	2760
DB	2701	AACAAAACTTCTGACACAAAGTATGCCCCATTTTGAAGAAATGGGCATGTTGTAGAA	2760
QY	2761	ACAGACATGGAGCTCTTAGTGTGCACTTTTCTGCAATTAATTCAGCGCCACCTGTAAGAG	2820
DB	2761	ACAGACATGGAGCTCTTAGTGTGCACTTTTCTGCAATTAATTCAGCGCCACCTGTAAGAG	2820
QY	2821	CTTACAGCTGATTGAGGGCAGCAGACAGATCAACATGGAGCCCAACCATGTTAGTCTG	2880
DB	2821	CTTACAGCTGATTGAGGGCAGCAGCAGACAGATCAACATGGAGCCCAACCATGTTAGTCTG	2880
QY	2881	TTGAGTGGTCCAGTACAGATGCTTATGGAGATTTCTTCTCGTCTCTGAGCTC	2940
DB	2881	TTGAGTGGTCCAGTACAGATGCTTATGGAGATTTCTTCTCGTCTCTGAGCTC	2940
QY	2941	ACCAGAACTCTGAAGGAGCTGAGACCTTAAGTGAAACTCTGCTGAGCACTCTGAGTGAAG	3000
DB	2941	ACCAGAACTCTGAAGGAGCTGAGACCTTAAGTGAAACTCTGCTGAGCACTCTGAGTGAAG	3000
QY	3001	AGTCTTTGTAAGACCTCTGAGACCCCTCTGCTGCTCTGAGACCTCTGCTGCTGCTG	3060
DB	3001	AGTCTTTGTAAGACCTCTGAGACCCCTCTGCTGCTCTGAGACCTCTGCTGCTGCTG	3060
QY	3061	TGTGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGCTGAGAGCCCAACAGACC	3120
DB	3061	TGTGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGCTGAGAGCCCAACAGACC	3120
QY	3121	CTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCCAACCTTTGCT	3180
DB	3121	CTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCCAACCTTTGCT	3180
QY	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCTGCTGCTGCTGCTGCTG	3240
DB	3181	CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCTGCTGCTGCTGCTGCTG	3240
QY	3241	ACGTCTGACTGCTGCAAGACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3300
DB	3241	ACGTCTGACTGCTGCAAGACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3300
QY	3301	CTAGACCTGAGCAGCAAACTGATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3360
DB	3301	CTAGACCTGAGCAGCAAACTGATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3360
QY	3361	AGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
DB	3361	AGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3420
QY	3421	ATGAGGAGGAACTGAGGCTCTGAGGAGGAGAACTTCACTGCTGCTGCTGCTGCTG	3480

Db 3421 ATGAGGACGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480
 Qy 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGCCCTGGATACGGGAGAGATGATATAGC 3540
 Db 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGGCCCTGGATACGGGAGAGATGATATAGC 3540
 Qy 3541 ACATCTCTCACTCAAGCGGAGAGACTCGGATCAGAGAGGGCGCTTCCCATCTGTGCTCAG 3600
 Db 3541 ACATCTCTCACTCAAGCGGAGAGACTCGGATCAGAGAGGGCGCTTCCCATCTGTGCTCAG 3600
 Qy 3601 GCTAATCTCAAACTCTGAGCGTGAAGAACTTCCCAATTCGTGAGATTCGAGAGGAA 3660
 Db 3601 GCTAATCTCAAACTCTGAGCGTGAAGAACTTCCCAATTCGTGAGATTCGAGAGGAA 3660
 Qy 3661 AGCTCCCAAGAGTGTACCGGTGGAACCTTGTGGTGGCTTCTCCCTGCTCAAGG 3720
 Db 3661 AGCTCCCAAGAGTGTACCGGTGGAACCTTGTGGTGGCTTCTCCCTGCTCAAGG 3720
 Qy 3721 GACCTGCATACGAAGCTTTGGGGAAGTACGATGACTTCTGGGGCCCAACGGGGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCTTTGGGGAAGTACGATGACTTCTGGGGCCCAACGGGGCTGTG 3780
 Qy 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840
 Db 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840
 Qy 3841 TCCTACCGCTGCCCAACACAGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Db 3841 TCCTACCGCTGCCCAACACAGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Qy 3901 ATTGAATCTGTGTGGGACCACTTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCACTTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Qy 3961 GTGGAGGGCCCTCTGTGACATCAAGGCTGAGCTTGGAGCTGTGGAAGCTGTGACCTTC 4020
 Db 3961 GTGGAGGGCCCTCTGTGACATCAAGGCTGAGCTTGGAGCTGTGGAAGCTGTGACCTTC 4020
 Qy 4021 CCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGCCAC 4080
 Db 4021 CCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCTGTTCCAAATGCCAC 4080
 Qy 4081 TTTAAAGAGAGGGGATGCTCTCTGGAGAGCCAGCCAGGGTGGAGCTGCATCACATAGTT 4140
 Db 4081 TTTAAAGAGAGGGGATGCTCTCTGGAGAGCCAGCCAGGGTGGAGCTGCATCACATAGTT 4140
 Qy 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTCTGAAATGATCAATATGCCCTG 4200
 Db 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGAGTCTCTCTGAAATGATCAATATGCCCTG 4200
 Qy 4201 CGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTTACCACGGGTCCATCCTCAGGAGTC 4260
 Db 4201 CGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTTACCACGGGTCCATCCTCAGGAGTC 4260
 Qy 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTCTCATTCGGAAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTCTCATTCGGAAGAACTGGAGCTCTGC 4320
 Qy 4321 TATCGAAGCCCTGGAGAACACAGCTGTTCTCGGAGTCTACGTTGGCCACTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAACACAGCTGTTCTCGGAGTCTACGTTGGCCACTTGGGATCA 4380
 Qy 4381 GGGATCAGGCTCGAAGTGAAGACAGAAAGATGAGACTCTGGTGTGGAGCCCTTGGT 4440
 Db 4381 GGGATCAGGCTCGAAGTGAAGACAGAAAGATGAGACTCTGGTGTGGAGCCCTTGGT 4440
 Qy 4441 AAACACGAGAGATCTCATGCTGCTCAACTCTCTGATCCCTCCAGCCCGCATAGCCGTACCT 4500
 Db 4441 AAACACGAGAGATCTCATGCTGCTCAACTCTCTGATCCCTCCAGCCCGCATAGCCGTACCT 4500
 Qy 4501 TCACCTCTGGATGCCCCCGAGTTGTGCACTTTGTGGACAGTATCGAGAGCAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCCCGAGTTGTGCACTTTGTGGACAGTATCGAGAGCAGCTGATA 4560

Qy 4561 GCCCAGTGACATCGGTGGAGGTTGTTCTTGGACAAACTGTCATGACAGGTGCTGAGCCAG 4620
 Db 4561 GCCCAGTGACATCGGTGGAGGTTGTTCTTGGACAAACTGTCATGACAGGTGCTGAGCCAG 4620
 Qy 4621 GAGCAGTACGAGAGGGTCTGGCTGAGAAACACGAGGCCACCGCAGATCGGAAGCTGTTTC 4680
 Db 4621 GAGCAGTACGAGAGGGTCTGGCTGAGAAACACGAGGCCACCGCAGATCGGAAGCTGTTTC 4680
 Qy 4681 AGCTTGAGCCAGTCTGGACCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4740
 Db 4681 AGCTTGAGCCAGTCTGGACCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4740
 Qy 4741 ACCCATCTCTCACTCATTTATGGAACCTCTGGAGAGAGGAGCAGCAAAAAGGAGCTCTCTGCCA 4800
 Db 4741 ACCCATCTCTCACTCATTTATGGAACCTCTGGAGAGAGGAGCAGCAAAAAGGAGCTCTCTGCCA 4800
 Qy 4801 CTCAGCAGTGAAGTATCAACACAGAGCCCTTGAACCTTTGAGTCTCTGGCTTTGGCTGACCC 4860
 Db 4801 CTCAGCAGTGAAGTATCAACACAGAGCCCTTGAACCTTTGAGTCTCTGGCTTTGGCTGACCC 4860
 Qy 4861 TTCTTTGGGTCTCAGTTTCTTCTCTCAAAACAAAGTTGCCATCTGGTTTGGCTTCCAGCA 4920
 Db 4861 TTCTTTGGGTCTCAGTTTCTTCTCTCAAAACAAAGTTGCCATCTGGTTTGGCTTCCAGCA 4920
 Qy 4921 CTAAAGTAAATGGAACCTTTGATGATGCCCTTCTGCTGGCAATTATGTCTCATGCCAGGATG 4980
 Db 4921 CTAAAGTAAATGGAACCTTTGATGATGCCCTTCTGCTGGCAATTATGTCTCATGCCAGGATG 4980
 Qy 4981 CCACAGGGGGCCCCAGTCCAGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Db 4981 CCACAGGGGGCCCCAGTCCAGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
 Qy 5041 GGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCCCAAGCCTAGA 5100
 Db 5041 GGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCCCAAGCCTAGA 5100
 Qy 5101 GCCCTCCGGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCATTTTGCCTC 5160
 Db 5101 GCCCTCCGGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCATTTTGCCTC 5160
 Qy 5161 TGGCTGTCTCAGGGTGAAGCCCAAAATTTGGGTTTACGCTGGAGGCCACGTTGATTC 5220
 Db 5161 TGGCTGTCTCAGGGTGAAGCCCAAAATTTGGGTTTACGCTGGAGGCCACGTTGATTC 5220
 Qy 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCACAGAGTAAAGTGAAGGAGTTT 5280
 Db 5221 TTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCACAGAGTAAAGTGAAGGAGTTT 5280
 Qy 5281 ATTCAAGAAATAAAGGAGTATCAGAGCTCTTTTAAAGTTTCTAGCAGGCTTTTCCAGTT 5340
 Db 5281 ATTCAAGAAATAAAGGAGTATCAGAGCTCTTTTAAAGTTTCTAGCAGGCTTTTCCAGTT 5340
 Qy 5341 TTTACAGAAAACCCCTATAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
 Db 5341 TTTACAGAAAACCCCTATAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
 Qy 5401 AAAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAAT 5444
 Db 5401 AAAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAAT 5444

RESULT 2
 AF310105 5404 bp mRNA linear PRI 04-MAY-2001
 LOCUS Homo sapiens NALP1 mRNA, complete cds.
 DEFINITION AF310105
 ACCESSION AF310105
 VERSION AF310105.1 GI:11096302
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1 (bases 1 to 5404).	Db	241	CCTGCTGGCCCTCGAGACCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGTGCT	300
AUTHORS	Martinon,F., Hofmann,K. and Tschoopp,J.	Qy	305	TGCAGGCCCTCTGGGCTCTCCCTCCCTCCCTGCTTTCTTCTACCACTCCCTCTATCGGC	364
TITLE	The pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation	Db	301	TGCAGGCCCTCTGGGCTCTCCCTCCCTCCCTGCTTTCTTCTACCACTCCCTCTATCGGC	360
JOURNAL	Curr. Biol. 11 (4), R118-R120 (2001)	Qy	365	GTCTATCTGTAGTGGCTCTGGATTTTATAAACTCGGTTCCGAATGCTGAATAAGAGCG	424
MEDLINE	21148093	Db	361	GTCTATCTGTAGTGGCTCTGGATTTTATAAACTCGGTTCCGAATGCTGAATAAGAGCG	420
PUBMED	11250163	Qy	425	GTAAGAGCAAGGCAAGGAG	484
REFERENCE	2 (bases 1 to 5404)	Db	421	GTAAGAGCAAGGCAAGGAG	480
AUTHORS	Martinon,F., Hofmann,K. and Tschoopp,J.	Qy	485	GAACATCCCCAGAGACACCTCTTAACTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAG	544
TITLE	Direct Submission	Db	481	GAACATCCCCAGAGACACCTCTTAACTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAG	540
JOURNAL	Submitted (28-SEP-2000) Institute of Biochemistry, University of Lausanne, Ch des Roveresses 155, Epalinges 1066, Switzerland	Qy	545	GCCTGGCTGTTCCTTGGAGTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	604
FEATURES	Location/Qualifiers	Db	541	GCCTGGCTGTTCCTTGGAGTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	600
source	1. 5404	Qy	605	TCGCCAATAAAGCGCACTCCAGGAGCTCTTGGGTTGAGACACCCCTCAGCCAGAGAGA	664
	/organism="Homo sapiens"	Db	601	TCGCCAATAAAGCGCACTCCAGGAGCTCTTGGGTTGAGACACCCCTCAGCCAGAGAGA	660
	/mol type="mRNA"	Qy	665	CGAGTGGCATGAGGTGGCTCTCGTACCTCGGTGGCTCAGTATGGGAGAGAGAGAGAGAG	724
	/db xref="taxon:9606"	Db	661	CGAGTGGCATGAGGTGGCTCTCGTACCTCGGTGGCTCAGTATGGGAGAGAGAGAGAGAG	720
	519..4808	Qy	725	ACCTAGCCCTCATACCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	784
	/notes="contains NACHT, LRR and PYD domains"	Db	721	ACCTAGCCCTCATACCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	780
	/codon_start=1	Qy	785	AGGGGAG	844
	/product="NALP1"	Db	781	AGGGGAG	840
	/protein_id="AAG30288.1"	Qy	845	CCAGCAACCCACCTCCACCGCAGTGTAAATGCCCTGATCCATGAATGCCCGGGGGT	904
	/db xref="GI:11096303"	Db	841	CCAGCAACCCACCTCCACCGCAGTGTAAATGCCCTGATCCATGAATGCCCGGGGGT	900
	/translation="MAGGAWGRACLVLEFLKKEELKEFOLLANKAHSRSSGTEPAQ	Qy	905	GCACCCAGGGCTCAG	964
	PKTSMEVASLYVAQYGRWDALHTWQMKLSLCAQAGAGAGAGAGAGAGAGAGAGAGAGAG	Db	901	GCACCCAGGGCTCAG	960
	PHLGSPTSQSTAVLMPWHELPAGCTGSRVRLQPLDPTSGRWRREISASLLYQAL	Qy	965	GGAGAGAAATCTCTGCTCTCCTCTACCAAGCTCTTCCAGAGCTCCCGAGAGAGAGAG	1024
	PSSPDHESQSPNAPTSTAVLGSWSPQSLAPREOAPGTOWPDETSGIYYTE	Db	961	GGAGAGAAATCTCTGCTCTCCTCTACCAAGCTCTTCCAGAGCTCCCGAGAGAGAGAG	1020
	IREREREKSEKRPWAAVVGTPPOAHTSLOPHHPWPBSVRESLCSSTWPKNEDFNO	Qy	1025	CTCCAGGAG	1084
	KETQLLLORPHRPSDPLVRSWPDYVEENRHLIEIRDLFGPLDQEPRLVILQG	Db	1021	CTCCAGGAG	1080
	AGIGKSTUARQVKEAWRGQYIDRFQFYFSCRELAQSKVSLAEILGKDGATP	Qy	1085	CCCCACCTCAGCCAGCCCTAGCACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1144
	APIRQLTSRPERLLFDLGVDFWVLEQPSSELCHWSQPPADALGSLGKTLIP	Db	1081	CCCCACCTCAGCCAGCCCTAGCACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1140
	EASFLITARTLQNLIPSLQARWVLEFSESRKEYFYFTDERQAIAPRLVK	Qy	1145	TGGATGAAACGTCAGAGAAATTTACTACAGAGAAATCAGAGAGAGAGAGAGAGAGAGAG	1204
	SNKELMALCLPWPWSLACTCLMQMKRKEKLTSTKTTTLLHYLAOLAOAPLPG	Db	1141	TGGATGAAACGTCAGAGAAATTTACTACAGAGAAATCAGAGAGAGAGAGAGAGAGAGAG	1200
	QLRDLCSLAEGWOKKTLFSPDRLKHLGDGAILSTPLKMGILQEHPIPLSVFIHL	Qy	1205	AGAAAGGAG	1264
	CQEPFAMSYVLEDKGRKHSNCIIDLKLEAYHGLPGASTFRLLGLLSDG	Db	1201	AGAAAGGAG	1260
	EREMENIFRLLSQGNLMQWPSLQQLQPHSLHSLCLYETRNKPLFUTQVMAHFE	Qy	1265	TACAGCCCCCACCACCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1324
	ILFVLKRVNLTLELDSNLSHAGSKVLSKTLRPRCLLETLAGLGLTAEBCDKL	Db	1261	TACAGCCCCCACCACCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
	LAFLGRANOTLKEFLDSNLSHAGSKVLSKTLRPRCLLETLAGLGLTAEBCDKL	Qy	1325	CCTGGAATAATGAGGATTTTAAACCAAAATTTACACAGCTGTCTTCTTACAAAGACCTC	1384
	ASVLSASPSLKELDLQNNLDVGVLELCEGRHLPACKLIRGLDQTLTSDMRBELR	Db	1321	CCTGGAATAATGAGGATTTTAAACCAAAATTTACACAGCTGTCTTCTTACAAAGACCTC	1380
	ALOEKRPQLLIPSRKRPSVMTPEGLDTGEMSNSTSLKRLGSRERAASHVAQNLK				
	LDVSKIFPIAEIAEESSEVVFVELLCVSPASQDLHKLPLDIDDFWGTGPVAT				
	EYVDKKNLYRHVFVAGSVWNTGLCFVMEAVTEIEFCVWQFQLEINPQHSWM				
	VAGPLDIDIAEPCGAEVAVHLPFHVALQGGHVDTSFQMAHFKEEGMLKPARVELHH				
	IVLENPSPGLVLMHINLALFVPTSVLLYHRVPEEVTFLYLLPSPDCSIRKE				
	LELCVPSGEDOLPFSFVYHGLSGIRLOVKDKQETLYVEALVKGDLMPATTLIP				
	ARIAVPSDIDAQLLHFVDQYREQLIARTVSEVLDLKLHGVLSQEQYERVAENR				
	PSOMKLFSLISOSWDRCKDGLYQALUKETHPHLMELWEKSKKGLLPUS"				
BASE COUNT	1284 a 1518 c 1467 g 1135 t				
ORIGIN					
Query Match	99.2%; Score 5400.8; DB 9; Length 5404;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 5402; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
Qy	5 CAGGGCTCGAGAGGCTGAGAGAACTGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG				
Db	1 CAGGGCTCGAGAGGCTGAGAGAACTGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAG				
Qy	65 TCTGAAAGCCCATTCCTGCTCTGCGGCTCTCCCAACCCCACTCTTCTCAGCCTTGACG				
Db	61 TCTGAAAGCCCATTCCTGCTCTGCGGCTCTCCCAACCCCACTCTTCTCAGCCTTGACG				
Qy	125 CTCAAGGTTGATCTCAGGAGTCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG				
Db	121 CTCAAGGTTGATCTCAGGAGTCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG				
Qy	185 ACAGTGAGGCTTCCACACACCCCATCTCCCGTCCACACATCTCCCTCAGCCTCACCTC				
Db	181 ACAGTGAGGCTTCCACACACCCCATCTCCCGTCCACACATCTCCCTCAGCCTCACCTC				
Qy	245 CCTGCTGGCCCTGGAGCCCATCCAGGAGCTCCCTATCAGCTGACTTCTTCCAGTGTCT				

QY	1385	ACCCAGAAAGCCAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTGGAGGAAATC	1444
DB	1381	ACCCAGAAAGCCAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTGGAGGAAATC	1440
QY	1445	GAGGACATTTAATGTAGATCAGAGACTTATTTGGCCCCAGCCCTGGATACCCAGAAACCTC	1504
DB	1441	GAGGACATTTAATGTAGATCAGAGACTTATTTGGCCCCAGCCCTGGATACCCAGAAACCTC	1500
QY	1505	GCATAGTCATCTGCAGGGGGCTGCTGGAATTTGGGAAGTCAACACTGGCCAGGCAAGTGA	1564
DB	1501	GCATAGTCATCTGCAGGGGGCTGCTGGAATTTGGGAAGTCAACACTGGCCAGGCAAGTGA	1560
QY	1565	AGGAAGCCTGGGGAGAGGCCAGCTGTATGGGAGCCGCTTCCAGCATGCTTCTACTTCA	1624
DB	1561	AGGAAGCCTGGGGAGAGGCCAGCTGTATGGGAGCCGCTTCCAGCATGCTTCTACTTCA	1620
QY	1625	GCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGAATCTCGCTGAGCTCATCGGAAAGATG	1684
DB	1621	GCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGAATCTCGCTGAGCTCATCGGAAAGATG	1680
QY	1685	GGACAGCCACTCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTTCA	1744
DB	1681	GGACAGCCACTCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTTCA	1740
QY	1745	TCCTCGATGGTGTAGATGAGCCAGGATGGTCTTGCAGAGCGGAGTCTGAGCTCTGTC	1804
DB	1741	TCCTCGATGGTGTAGATGAGCCAGGATGGTCTTGCAGAGCGGAGTCTGAGCTCTGTC	1800
QY	1805	TGCATCTGGAGCCAGCCAGCCGCGGATGCACTGCTGGGCAATTTGCTGGGGAAAACTA	1864
DB	1801	TGCATCTGGAGCCAGCCAGCCGCGGATGCACTGCTGGGCAATTTGCTGGGGAAAACTA	1860
QY	1865	TACTTCCGAGGATCCTTCTGATCAGCGCTCGACCAAGCTCTGCAGAACCTCATTC	1924
DB	1861	TACTTCCGAGGATCCTTCTGATCAGCGCTCGACCAAGCTCTGCAGAACCTCATTC	1920
QY	1925	CTTCTTTGGAGGAGGACGCTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGGAAGG	1984
DB	1921	CTTCTTTGGAGGAGGACGCTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGGAAGG	1980
QY	1985	AATATTTCTACAGATATTTACAGATGAAAGCAAGCAATTAAGACCTTTAGTTGGTCA	2044
DB	1981	AATATTTCTACAGATATTTACAGATGAAAGCAAGCAATTAAGACCTTTAGTTGGTCA	2040
QY	2045	AATCAAAACAAAGCTCTGGGCCCTGTCTTGTGCCCCCTGGGTGCTCGCTGGCCCTGCA	2104
DB	2041	AATCAAAACAAAGCTCTGGGCCCTGTCTTGTGCCCCCTGGGTGCTCGCTGGCCCTGCA	2100
QY	2105	CTTGGCTGATGCAGAGATGAAGCGAAGGAAAACTCACTGACTTCCAAAGACCACCA	2164
DB	2101	CTTGGCTGATGCAGAGATGAAGCGAAGGAAAACTCACTGACTTCCAAAGACCACCA	2160
QY	2165	CAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGCCATTTGGACCCACG	2224
DB	2161	CAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGCCATTTGGACCCACG	2220
QY	2225	TCAGAGCCTCTGCTCTGTGGTCTGAGGGGATCTGGCAAAAAAGACCCCTTTTCAGTC	2284
DB	2221	TCAGAGCCTCTGCTCTGTGGTCTGAGGGGATCTGGCAAAAAAGACCCCTTTTCAGTC	2280
QY	2285	CAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATCATCTCCACTTCTTGAAGATGG	2344
DB	2281	CAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATCATCTCCACTTCTTGAAGATGG	2340
QY	2345	GTATTTCTCAAGAGCAGCCCATCCCTCTGAGCTACAGCTTCAATCACTCTGTTTCCAAAG	2404
DB	2341	GTATTTCTCAAGAGCAGCCCATCCCTCTGAGCTACAGCTTCAATCACTCTGTTTCCAAAG	2400
QY	2405	AGTTCTTTGAGCAATGTCTTGTCTTGGAGATGAGAGGGGAGAGTAACATTCTA	2464
DB	2401	AGTTCTTTGAGCAATGTCTTGTCTTGGAGATGAGAGGGGAGAGTAACATTCTA	2460

QY	2465	ATTGCATCATAGATTTTGAAAGAGCGCTAGAGCATATGGAATACATGGCCTGTTTGGGG	2524
DB	2461	ATTGCATCATAGATTTTGAAAGAGCGCTAGAGCATATGGAATACATGGCCTGTTTGGGG	2520
QY	2525	CATCAACACACGTTTCTATTGGGCTGTTAAAGTGTAGGGGGAGAGAGATGGAGA	2584
DB	2521	CATCAACACACGTTTCTATTGGGCTGTTAAAGTGTAGGGGGAGAGAGATGGAGA	2580
QY	2585	ACATCTTTCACTGCGGCTGTCTCAGGGGAGAACTCTGATCAGTGGGTCCCGTCCCTGC	2644
DB	2581	ACATCTTTCACTGCGGCTGTCTCAGGGGAGAACTCTGATCAGTGGGTCCCGTCCCTGC	2640
QY	2645	AGTGTCTCTGAGCCACACTCTCTGGAGTCCCTCACTGCTGTGTACAGACTCGGAACA	2704
DB	2641	AGTGTCTCTGAGCCACACTCTCTGGAGTCCCTCACTGCTGTGTACAGACTCGGAACA	2700
QY	2705	AAACGTTCTTCTGACACAGTGTATGCCCCATTTTCGAGAAATGGGCATGTGTAGAAAACAG	2764
DB	2701	AAACGTTCTTCTGACACAGTGTATGCCCCATTTTCGAGAAATGGGCATGTGTAGAAAACAG	2760
QY	2765	ACATGGAGCTCTTAGTGTGCACTTTCTGCATTTAAATTCAGCGGCCACCTGGAAGAGCTTC	2824
DB	2761	ACATGGAGCTCTTAGTGTGCACTTTCTGCATTTAAATTCAGCGGCCACCTGGAAGAGCTTC	2820
QY	2825	AGTGAATTTGAGGGCAGGAGCAGACAGATCAACATGAGAGCCCCACCATGTGTCTCTGTCA	2884
DB	2821	AGTGAATTTGAGGGCAGGAGCAGACAGATCAACATGAGAGCCCCACCATGTGTCTCTGTCA	2880
QY	2885	GGTGGTCCAGTACAGATGCTTATGGCAGATTTCTTCTCGTCTCAAGGTCAACA	2944
DB	2881	GGTGGTCCAGTACAGATGCTTATGGCAGATTTCTTCTCGTCTCAAGGTCAACA	2940
QY	2945	GAACCTCAAGAGCTGACCTAAGTGAAGAACTCGCTGAGCCACTCTCAGTGAAGAGTC	3004
DB	2941	GAACCTCAAGAGCTGACCTAAGTGAAGAACTCGCTGAGCCACTCTCAGTGAAGAGTC	3000
QY	3005	TTTGTGAAGACCTTGAGACGCTCTGCTGCTCTTGAGAGACCTTGGGTGCTGGCTGTG	3064
DB	3001	TTTGTGAAGACCTTGAGACGCTCTGCTGCTCTTGAGAGACCTTGGGTGCTGGCTGTG	3060
QY	3065	GCCTCAGAGCTGAGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCACAGACCTTGA	3124
DB	3061	GCCTCAGAGCTGAGACTGCAAGGACCTTGGCTTTGGGCTGAGAGCCACAGACCTTGA	3120
QY	3125	CGAGCTCGACTGAGCTTCAATGTCTCAGGATGCTGGAGGCCAAACCTTTGCCAGA	3184
DB	3121	CGAGCTCGACTGAGCTTCAATGTCTCAGGATGCTGGAGGCCAAACCTTTGCCAGA	3180
QY	3185	GACTGAGACAGCCGAGCTGCAAGCTTACAGCGACTGACGCTGGTGGCTCAGCTCAGT	3244
DB	3181	GACTGAGACAGCCGAGCTGCAAGCTTACAGCGACTGACGCTGGTGGCTCAGCTCAGT	3240
QY	3245	CTGACTGTGCGAGGACCTGGCTCTGTCTTAGTCCAGCCCCAGCCCTGGAAGAGCTAG	3304
DB	3241	CTGACTGTGCGAGGACCTGGCTCTGTCTTAGTCCAGCCCCAGCCCTGGAAGAGCTAG	3300
QY	3305	ACCTCAGCAGAAACCTGATGACGTTGGCTGCGACTGCTGTGAGGGGCTCAGCG	3364
DB	3301	ACCTCAGCAGAAACCTGATGACGTTGGCTGCGACTGCTGTGAGGGGCTCAGCG	3360
QY	3365	ATCCTGCTGCAAACTCATACGCTGGGGCTGGACCACTCTGAGTGATGATGATGA	3424
DB	3361	ATCCTGCTGCAAACTCATACGCTGGGGCTGGACCACTCTGAGTGATGATGATGA	3420
QY	3425	GGCAGGAACTGAGGGCCCTGGAGGAGGAGAAACCTCAGCTGCTCATCTTCAGCAGCCGA	3484
DB	3421	GGCAGGAACTGAGGGCCCTGGAGGAGGAGAAACCTCAGCTGCTCATCTTCAGCAGCCGA	3480
QY	3485	AACCAAGTGTATGATGACCCCTACTGAGGCTGGATACCGGAGAGATGATGATGATGAT	3544
DB	3481	AACCAAGTGTATGATGACCCCTACTGAGGCTGGATACCGGAGAGATGATGATGATGAT	3540
QY	3545	CCTCACTCAAGCGGAGAGACTCGGATCAGAGAGGGCGGCTTCCCATGTGTGCTCAGGCTA	3604

Db 3541 ||||| CCTCACTAAGCGGACAGACTCGGATCAGAGAGCGCGCTTCCCATGTTGCTCAGGCTA 3600
Qy 3605 ATCTCAAACTCTGGAGCTGAGCAAGATCTTCCAAATTCGTGAGATTGCGAGGAAAGCT 3664
Db 3601 ATCTCAAACTCTGACGCTGAGCAAGATCTTCCAAATTCGTGAGATTGCGAGGAAAGCT 3660
Qy 3665 CCCCAGAGTAGTACCGGTGGAACTCTTGTGCGTGCTTCTCTGCTCTCAAGGGGACC 3724
Db 3661 CCCCAGAGTAGTACCGGTGGAACTCTTGTGCGTGCTTCTCTGCTCTCAAGGGGACC 3720
Qy 3725 TGCATACGAAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCCACGGGGCTGTGGCTA 3784
Db 3721 TGCATACGAAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCCACGGGGCTGTGGCTA 3780
Qy 3785 CTGAGGTAGTTGACAAAGAAAGAACTCTGTCGAGTTTCACTTCCCTGTAGCTGGCTCT 3844
Db 3781 CTGAGGTAGTTGACAAAGAAAGAACTCTGTCGAGTTTCACTTCCCTGTAGCTGGCTCT 3840
Qy 3845 ACCGTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAGAAGCGGTGACCGTTGAGATTG 3904
Db 3841 ACCGTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAGAAGCGGTGACCGTTGAGATTG 3900
Qy 3905 AATTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGCACAGCTGGATGGTG 3964
Db 3901 AATTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGCACAGCTGGATGGTG 3960
Qy 3965 CAGGGCTCTGTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACTTCCCTC 4024
Db 3961 CAGGGCTCTGTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACTTCCCTC 4020
Qy 4025 ACTTTGTGGCTCTCCAAGGGGGCCATGTGGACACATCCCTGTTTCCAAATGGGCCACTT 4084
Db 4021 ACTTTGTGGCTCTCCAAGGGGGCCATGTGGACACATCCCTGTTTCCAAATGGGCCACTT 4080
Qy 4085 AGAGAGGGGATGCTCTGGAGAACCGACGAGGTGAGCTGCAATGATTCATGTTCTGG 4144
Db 4081 AGAGAGGGGATGCTCTGGAGAACCGACGAGGTGAGCTGCAATGATTCATGTTCTGG 4140
Qy 4145 AAAAAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAATGATCCATAATGCCCTGCGCT 4204
Db 4141 AAAAAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAATGATCCATAATGCCCTGCGCT 4200
Qy 4205 TCATTCCTGCTCAGCTCTGTGGTGTGCTTTTACACCGCGTCCATCTCGAGGAAGTCACT 4264
Db 4201 TCATTCCTGCTCAGCTCTGTGGTGTGCTTTTACACCGCGTCCATCTCGAGGAAGTCACT 4260
Qy 4265 TCACCTCTACCTGATCCCAAGTGA CTGCTCCATTCGGAAGAACTGGAGCTCTGTATC 4324
Db 4261 TCACCTCTACCTGATCCCAAGTGA CTGCTCCATTCGGAAGAACTGGAGCTCTGTATC 4320
Qy 4325 GAAGCCTGGAGAGACCACTCTCTCGAGTTCTACGTTGGCCACTTGGGATCAGGGA 4384
Db 4321 GAAGCCTGGAGAGACCACTCTCTCGAGTTCTACGTTGGCCACTTGGGATCAGGGA 4380
Qy 4385 TCAGGCTGCAAGTGAAGACAAGAAAGATGAGACTCTGTGTGGGAGGCTTGGTGAAC 4444
Db 4381 TCAGGCTGCAAGTGAAGACAAGAAAGATGAGACTCTGTGTGGGAGGCTTGGTGAAC 4440
Qy 4445 CAGGAGATCTATGCTGCAACTACTCTGTATTCCTCAGCCCGCATAGCGCTACCTTAC 4504
Db 4441 CAGGAGATCTATGCTGCAACTACTCTGTATTCCTCAGCCCGCATAGCGCTACCTTAC 4500
Qy 4505 CTCTGGATGCCCGCAGTCTCTGCACTTGTGGACCAAGTATCAGAGCAGCTGATAGCCC 4564
Db 4501 CTCTGGATGCCCGCAGTCTCTGCACTTGTGGACCAAGTATCAGAGCAGCTGATAGCCC 4560
Qy 4565 GAGTGACATCGGTGGAGTTGTCTTGGACAACTGTGATGACAGAGTGTGAGCAGAGC 4624
Db 4561 GAGTGACATCGGTGGAGTTGTCTTGGACAACTGTGATGACAGAGTGTGAGCAGAGC 4620
Qy 4625 AGTACGAGAGGTGCTGGCTGAGAACACAGAGGCCCGACGAGCTGCGGAAGCTTTCAGCT 4684

Db 4621 AGTACGAGAGGTGCTGGCTGAGAACACAGAGCCCGACGATGGGAAGCTGTTCACT 4680
Qy 4685 TGAGCCAGTCTCTGGACCGGAAGTGCAAAGATGGAATCTTCAAGCCCTGGAAGGAGACC 4744
Db 4681 TGAGCCAGTCTCTGGACCGGAAGTGCAAAGATGGAATCTTCAAGCCCTGGAAGGAGACC 4740
Qy 4745 ATCTCACCTATTTATGAACTCTGGGAGAAAGGAGCAAAAAGGACTCTCTGCCACTCA 4804
Db 4741 ATCTCACCTATTTATGAACTCTGGGAGAAAGGAGCAAAAAGGACTCTCTGCCACTCA 4800
Qy 4805 GCAGCTGAAGTATCAACACAGCCCTTGCACCTTGCAGTCTCTGGCTTGGCTGACCTTCT 4864
Db 4801 GCAGCTGAAGTATCAACACAGCCCTTGCACCTTGCAGTCTCTGGCTTGGCTGACCTTCT 4860
Qy 4865 TTGGTCTCAGTCTTCTTCTCTGCAACAAGTTGCCATCTCGTCTTGGCTTCCAGACTAA 4924
Db 4861 TTGGTCTCAGTCTTCTTCTCTGCAACAAGTTGCCATCTCGTCTTGGCTTCCAGACTAA 4920
Qy 4925 AGTAATGGAATCTTGCATGATGCTTGTCTGGGCATTTATGTCTCATGCGGATGCCAC 4984
Db 4921 AGTAATGGAATCTTGCATGATGCTTGTCTGGGCATTTATGTCTCATGCGGATGCCAC 4980
Qy 4985 AGGGGGCCCCAGTCCAGGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5044
Db 4981 AGGGGGCCCCAGTCCAGGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5040
Qy 5045 AGACCCCTGCAGACCTCATAGAGCCTCATCTGTGTGGGCCACAGCAGCCAGCCTAGAGCCC 5104
Db 5041 AGACCCCTGCAGACCTCATAGAGCCTCATCTGTGTGGGCCACAGCAGCCAGCCTAGAGCCC 5100
Qy 5105 TCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTCTGGC 5164
Db 5101 TCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTCTGGC 5160
Qy 5165 TGTGTACAGGCTGAGCCCCAAAATTTGGGTTTTCAGCTGGGAGGCCAGTGGATTCTTGG 5224
Db 5161 TGTGTACAGGCTGAGCCCCAAAATTTGGGTTTTCAGCTGGGAGGCCAGTGGATTCTTGG 5220
Qy 5225 CTTTGTACAGGAAGATCTACAGAGCAAGCCCAAGAGTAAAGTGAAGGAAGTTTATTC 5284
Db 5221 CTTTGTACAGGAAGATCTACAGAGCAAGCCCAAGAGTAAAGTGAAGGAAGTTTATTC 5280
Qy 5285 AGAAATAAAGAGATACAGCTCTTTTAGAATTTGTCTAGCAGCTTTCAGTTTTTTA 5344
Db 5281 AGAAATAAAGAGATACAGCTCTTTTAGAATTTGTCTAGCAGCTTTCAGTTTTTTA 5340
Qy 5345 CCAGAAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTAAAGAAATTAATAATAACA 5404
Db 5341 CCAGAAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTAAAGAAATTAATAATAACA 5400
Qy 5405 AAAA 5408
Db 5401 AAAA 5404

RESULT 3
LOCUS AX459865
DEFINITION Sequence 8 from Patent WO0240668.
ACCESSION AX459865
VERSION AX459865.1 GI:21725641
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Tschoep, J. and Martinon, F.
TITLE Proteins and dna sequences underlying these proteins used for
treating inflammations
JOURNAL Patent: WO 0240668-A 8 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
Location/Qualifiers

Db 2133 AGCATATGGAATACATATGGGCTGTTGGGGATCAACCAACAGCTTTCCTATTTGGGCTGTT 2192
 QY 2556 AAGTATGAGGGGAGAGAGATGAGAAACATCTTTCACTGCGGCTGTCTCAGGGGAG 2615
 Db 2193 AAGTATGAGGGGAGAGAGATGAGAAACATCTTTCACTGCGGCTGTCTCAGGGGAG 2252
 QY 2616 GAACCTGATGAGTGGGTCCCGTCCCTGAGCTGCTGCGACGACATCTCTCTGAGTGC 2675
 Db 2253 GAACCTGATGAGTGGGTCCCGTCCCTGAGCTGCTGCGACGACATCTCTCTGAGTGC 2312
 QY 2676 CTTCCACTCTGTTAGAGACTCGGAACAAAGGTTCTTGACACAGTGCATGSCCATTT 2735
 Db 2313 CTTCCACTCTGTTAGAGACTCGGAACAAAGGTTCTTGACACAGTGCATGSCCATTT 2372
 QY 2736 CGAAGAAATGGGATGCTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCAT 2795
 Db 2373 CGNAGAAATGGGATGCTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCAT 2432
 QY 2796 TAAATTCAGCGCCACGTTGAAGCTTCAGCTGATTGAGGGCAGGACACAGATCAAC 2855
 Db 2433 TAAATTCAGCGCCACGTTGAAGCTTCAGCTGATTGAGGGCAGGACACAGATCAAC 2492
 QY 2856 ATGGAGCCCAACATGGTCTGTTCAAGTGGGTCCAGTCCAGTCCAGTGCATTTATGGCA 2915
 Db 2493 ATGGAGCCCAACATGGTCTGTTCAAGTGGGTCCAGTCCAGTCCAGTGCATTTATGGCA 2552
 QY 2916 GATTCCTTTCTCGTCTCTCAAGTCCACAGAACCTGAAGGAGCTGGACCTTAAGTGA 2975
 Db 2553 GATTCCTTTCTCGTCTCTCAAGTCCACAGAACCTGAAGGAGCTGGACCTTAAGTGA 2612
 QY 2976 CTGCTGAGCCACTCTGCTGAGTGAAGTCTTTGTAAAGCCTTGAGACGCCCTGCTGCT 3035
 Db 2613 CTGCTGAGCCACTCTGCTGAGTGAAGTCTTTGTAAAGCCTTGAGACGCCCTGCTGCT 2672
 QY 3036 CTTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3095
 Db 2673 CTTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2732
 QY 3096 CTTTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTTCAATGTGCTCAC 3155
 Db 2733 CTTTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTTCAATGTGCTCAC 2792
 QY 3156 GGATGCTGAGGCAAAACATTTTCCAGAGACTGAGACAGCCGAGTCAAGCTACAGCG 3215
 Db 2793 GGATGCTGAGGCAAAACATTTTCCAGAGACTGAGACAGCCGAGTCAAGCTACAGCG 2852
 QY 3216 ACTGAGCTGCTGAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3275
 Db 2853 ACTGAGCTGCTGAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2912
 QY 3276 TAGTGCAGCCAGCTGAGAGCTAGACCTGACAGCAACCTGGATGACCTTGG 3335
 Db 2913 TAGTGCAGCCAGCTGAGAGCTAGACCTGACAGCAACCTGGATGACCTTGG 2972
 QY 3336 CGTGCAGCTGCTGCTGAGGGCTCAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3395
 Db 2973 CGTGCAGCTGCTGCTGAGGGCTCAGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3032
 QY 3396 GGACCAAGAACTCTGATGATGAGTGAAGCAAGAACTGAGGGCCCTTGAGCAGAGAA 3455
 Db 3033 GGACCAAGAACTCTGATGATGAGTGAAGCAAGAACTGAGGGCCCTTGAGCAGAGAA 3092
 QY 3456 ACCTCAGCTGCTATCTTACAGCAGCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3515
 Db 3093 GCCTCAGCTGCTATCTTACAGCAGCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3152
 QY 3516 GGATAGGGAGAGATGATTAATAGCAGATCCTCACTCAAGCGCAGAGACTCGGATCAGA 3575
 Db 3153 GGATAGGGAGAGATGATTAATAGCAGATCCTCACTCAAGCGCAGAGACTCGGATCAGA 3212
 QY 3576 GAGGGGCTTCCATGCTGCTCAGGCTAATCTCAAACTCTCGGACGTCGAGCAAGATCTT 3635
 Db 3213 GAGGGGCTTCCATGCTGCTCAGGCTAATCTCAAACTCTCGGACGTCGAGCAAGATCTT 3272

QY 3636 CCCAATTCTGAGATTGAGAGGAAAGCTCCCAAGAGTAGTACCGGTGGAACTCTTGTG 3695
 Db 3273 CCCAATTCTGAGATTGAGAGGAAAGCTCCCAAGAGTAGTACCGGTGGAACTCTTGTG 3332
 QY 3696 CGTGCCTTCTCTGCTCTCAAGGGGACTGCATAGAGCCTTTTGGGACTGACGATGA 3755
 Db 3333 CGTGCCTTCTCTGCTCTCAAGGGGACTGCATAGAGCCTTTTGGGACTGACGATGA 3392
 QY 3756 CTTCTGGGGCCCAACCGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGT 3815
 Db 3393 CTTCTGGGGCCCAACCGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGT 3452
 QY 3816 CCGAGTTCACTTCCCTGTAGTGGCTCTTACCGCTGGGCCAAACACGGGTCTCTGCTTGT 3875
 Db 3453 CCGAGTTCACTTCCCTGTAGTGGCTCTTACCGCTGGGCCAAACACGGGTCTCTGCTTGT 3512
 QY 3876 GATGAGAGAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTTGGTGA 3935
 Db 3513 GGTGAGAGAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTTGGTGA 3572
 QY 3936 GATCAACCCACAGCAGCAGCTGGATGGTGGAGGCTCTGTGGACATCAAGGCTGAGCC 3995
 Db 3573 GATCAACCCACAGCAGCAGCTGGATGGTGGAGGCTCTGTGGACATCAAGGCTGAGCC 3632
 QY 3996 TGGAGCTGTGGAAGCTGTGCACCTCCCTCACCTTTTGTGGCTCTTCAAGGGGGCCATGTGA 4055
 Db 3633 TGGAGCGGTGGAAGCTGTGCACCTCCCTCACCTTTTGTGGCTCTTCAAGGGGGCCATGTGA 3692
 QY 4056 GACATCCCTGTTTCCAAATGGGCCCACTTTAAAGAGAGGGGATGCTCTTGGAGAGCCAGC 4115
 Db 3693 CACATCCCTGTTTCCAAATGGGCCCACTTTAAAGAGAGGGGATGCTCTTGGAGAGCCAGC 3752
 QY 4116 CAGGTGTGAGCTGCATCATAGTTCTGGAACCCAGCTTCTCCCTTGGGAGTCT 4175
 Db 3753 CAGGTGTGAGCTGCATCATAGTTCTGGAACCCAGCTTCTCCCTTGGGAGTCT 3812
 QY 4176 CTTGAAAATGATCCATAATGCCCTGCTGCTTCAATCCCGTCACTCTGTGGTGTGCTTTA 4235
 Db 3813 CTTGAAAATGATCCATAATGCCCTGCTGCTTCAATCCCGTCACTCTGTGGTGTGCTTTA 3872
 QY 4236 CCACCGCTCATCTCTGAGGAAAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTC 4295
 Db 3873 CCACCGCTCATCTCTGAGGAAAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTC 3932
 QY 4296 CATTGGAAGAACTGAGCTCTGCTATCGAGCCCTGAGAGAGCCAGCTGTCTCGGA 4355
 Db 3933 CATTGGAAGAACTGAGCTCTGCTATCGAGCCCTGAGAGAGCCAGCTGTCTCGGA 3992
 QY 4356 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGATGA 4415
 Db 3993 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGATGA 4052
 QY 4416 GACTGTGTGGGAGGCTTGGTGAACCAAGAGATCTCATGCTGCAACTCTCTGAT 4475
 Db 4053 GACTGTGTGGGAGGCTTGGTGAACCAAGAGATCTCATGCTGCAACTCTCTGAT 4112
 QY 4476 CCCTCAGCCCGCATAGCCGTACCTTCACTCTGATGCCCGCAGCTTCTGCACCTTCT 4535
 Db 4113 CCCTCAGCCCGCATAGCCGTACCTTCACTCTGATGCCCGCAGCTTCTGCACCTTCT 4172
 QY 4536 GGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGGTTGTCTTGGACAA 4595
 Db 4173 GGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGGTTGTCTTGGACAA 4232
 QY 4596 ACTGCATGCAAGGTGCTGAGCCAGAGCAGTACGAGAGGGTCTGGCTGAGAAACAGAG 4655
 Db 4233 ACTGCATGCAAGGTGCTGAGCCAGAGCAGTACGAGAGGGTCTGGCTGAGAAACAGAG 4292
 QY 4656 GCCCAGCAGATGCGGAAGCTGTTGAGCTTGAAGCAGTCTTGGGACCGGAGTGCAGAA 4715
 Db 4293 GCCCAGCAGATGCGGAAGCTGTTGAGCTTGAAGCAGTCTTGGGACCGGAGTGCAGAA 4352

QY	4716	TGGACTCTACCAAGCCTGAAGGAGACCCATCTCCTCACTCATATGAACTCTGGAGAA	4775
Db	4353	TGGACTCTACCAAGCCTGAAGGAGACCCATCTCCTCACTCATATGAACTCTGGAGAA	4412
QY	4776	GGCAGCAAAAGGAGACTCTGCGACTCAGAGCTGAAGATCAACACAGGCCCTTGACC	4835
Db	4413	GGCAGCAAAAGGAGACTCTGCGACTCAGAGCTGAAGATCAACACAGGCCCTTGACC	4472
QY	4836	CTTGAGTCTGGCTTTGGCTTGAGCCCTCTTTGGGTCTCAGTTTCTCTCTGCAACAG	4895
Db	4473	CTTGAGTCTGGCTTTGGCTTGAGCCCTCTTTGGGTCTCAGTTTCTCTCTGCAACAG	4532
QY	4896	TTGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTCTGCG	4955
Db	4533	TTGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTCTGCG	4592
QY	4956	GCATTATGTTCATGCCAGGAGTGCACAGGGGGCCCGAGTCCAGTGGCTTAACAGCA	5015
Db	4593	GCATTATGTTCATGCCAGGAGTGCACAGGGGGCCCGAGTCCAGTGGCTTAACAGCA	4652
QY	5016	TCTCAGGGAATGCCATCTGGAGCTGGCAGACCCCTGCAGACCTCATAGAGCTCATCT	5075
Db	4653	TCTCAGGGAATGCCATCTGGAGCTGGCAGACCCCTGCAGACCTCATAGAGCTCATCT	4712
QY	5076	GTGGGCCACAGCAAGCCTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGGAATA	5135
Db	4713	GTGGGCCACAGCAAGCCTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGGAATA	4772
QY	5136	GGAGGACATGGAACCATTTGCTCTGCTGTGTCTCAGGGTGAGCCCCCAAAATTTGGGT	5195
Db	4773	GGAGGACATGGAACCATTTGCTCTGCTGTGTCTCAGGGTGAGCCCCCAAAATTTGGGT	4832
QY	5196	TCAGCTGGGAGCCAGCTGGATCTTGCTTTGTACAGAGAGATCTACAGCAAGC	5255
Db	4833	TCAGCTGGGAGCCAGCTGGATCTTGCTTTGTACAGAGAGATCTACAGCAAGC	4892
QY	5256	ACAGAGTAAAGTGAAGGAAGTTTATCAGAAAAATAAGGAGTATCAGAGCTCTTTAG	5315
Db	4893	ACAGAGTAAAGTGAAGGAAGTTTATCAGAAAAATAAGGAGTATCAGAGCTCTTTAG	4952
QY	5316	AATTTCTAGCAGGCTTTCCAGTTTATACAGAAAAACCCCTATAAATTAATAATTTT	5375
Db	4953	AATTTCTAGCAGGCTTTCCAGTTTATACAGAAAAACCCCTATAAATTAATAATTTT	5012
QY	5376	ACTTAATTTAGAAATTAATAATAAATAAAGAAATAAATAAAGAAATAAGAA	5435
Db	5013	ACTTAATTTAGAAATTAATAATAAATAAAGAAATAAATAAAGAAATAAGAA	5072
QY	5436	GTTA 5439	
Db	5073	GTTA 5076	
RESULT 4			
LOCUS	AF298548		
DEFINITION	Homo sapiens caspase recruitment domain protein 7 mRNA, complete		
ACCESSION	AF298548		
VERSION	AF298548.1		
KEYWORDS	GI:10198208		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Bertin, J. and DiStefano, P.S.		
JOURNAL	The PYRIN domain: a novel motif found in apoptosis and inflammation		
MEDLINE	Proteins		
PUBMED	Cell Death Differ. 7 (12), 1273-1274 (2000)		
REFERENCE	2 (bases 1 to 5100)		

AUTHORS	Bertin, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-AUG-2000) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA		
FEATURES	Location/Qualifiers		
source	1..5100		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
CDS	160..4449		
	/note="CARD7; nucleotide-binding site/leucine-rich repeat family member (NBS/LRR); contains PYRIN and CARD domains"		
	/product="caspase recruitment domain protein 7"		
	/protein_id="AAG15254.1"		
	/db_xref="GI:10198208"		
	/translation="MAGAWGRGLACYLEFLKKEELKEPOLLANKAHRRSSSGETPAQ PEKTSMEVASYLVAQGEQAWDLALHTWEQMLRSLCAQAQAGAGHSFPFSPSE PHLGSPSOPTSTAVLMPWHELHPAGCTGSSRRVLRQLPDTSGRRWREISASHLYQAL PSSDPHSPQSPNAPTSSTAVLWGSWGPQPSLAPRQEPAPQWPLEDSTGYIYTE IREEREREKSRPRPAAVCTPQAHSLSQPHHPWEPVSRESLCTWPKNEDFNQ KFTQLLLQRPSPQDLVXKSPDPYVEENRHLIEIRDLFPGDLTQERVLITQ AAGIKSTLAKQVKEAMRGVIGYDFQHFYFSCRELAQKVKVSLAGLIGRDGTATP APIRQILSRPERLLFILGDVDEPGWLOPSELSLHWSQPQADALLGLGKTLIP EASFILITARTALQNLIPSLQPAQWVEVLGFSSESRKEYFYFYFDERQAIARFLVK SNKELMALCLVPWVSLACTCLMOQMKREKLTITSTFKWGLIQEHPILSYSIHL QRLDLSLAASGIWQKTLFSPDLRLKHGLDGAIISTFKWGLIQEHPILSYSIHL CFQFFFAAMSVLEDEKRGKHSNCIDLEKTLKLEAYGHGLFGASTTRFLGLSDEG EREMENTPHCLSQGNLMQWPSLQLLQPHLSLESLHCLYETRNKTLTQVMAHFEE MCMVETDEMLLVCTFCIKFSRHVKQLQIEGRHRSWSPMVVLFWMVPTADYQW ILFSLVKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLETURLACGLTAEDCKD LAFGLRANOTLUTELDLSFNVLMDSAKHLQRLQPSCKLQRLQSVCLGSLSDCCQDL ASVLSASPKELDLQONLDDVGRLLCEGLRHPACKLIRLGDQTLTSLDEMROELR ALBEKPOLLIIFSRKESVMTPIEGLDTGENSTSSSLKROLRSGERASVQAQNLK LLDVSKPIPIAIEAESSEPVVVELLCVPSASQDHLTKPLGTDHDFWGTGCVAT EVQDEKRLXRVHFVAGSYRMPNTGLCFVVRVATVIEFCVMDQFLGEINPQSHW VAGPLLDIKAEPPGAVEAHLPHFVALQGHVDTSLFQVAHFKEGMLLEKARVELHWH IVLNPSPLIGVLLKMIHNLRFIPVTSVLLVHRLHPEVTHLYLIPSDCSIRKE LELCVSPGEDQLSEFVYVGHGSLRLOVKKDETLVWEALVKPGDMLPATTLLIPP ACTAVPSPLDAPQLLHFVDYQREQLIARVTSVEVVDLKLHGQVLSQEQYERVLANTR PSQWRKLFSLQSQRDRCKDGLYQALKEHPHLMELMEKSKGLLPLSS"		
BASE COUNT	1251 a 1374 c 1406 g 1069 t		
ORIGIN			
Query Match	89.7%; Score 4882.4; DB 9; Length 5100;		
Best Local Similarity	99.5%; Pred. No. 0;		
Matches 4898; Conservative	0; Mismatches 26; Indels 0; Gaps 0;		
QY	516	GACAGAGATGGCTGGCGGAGCCTGGGGCCGCTGTCCTGCCAATAAAGCGCACTCCAGAGCTTTC	575
Db	153	GAACAAGATGGCTGGCGGAGCCTGGGGCCGCTGTCCTGCCAATAAAGCGCACTCCAGAGCTTTC	212
QY	576	GGAGGAGCTGAAGGATTCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTTTC	635
Db	213	GGAGGAGCTGAAGGATTCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTTTC	272
QY	636	GGGTGAGACACCCGCTCAGCAGAGACCGAGTGGCATGGAGTGGCTGCTGCTACTCTGCT	695
Db	273	GGGTGAGACACCCGCTCAGCAGAGACCGAGTGGCATGGAGTGGCTGCTGCTACTCTGCT	332
QY	696	GGCTCAGTATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGG	755
Db	333	GGCTCAGTATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGG	392
QY	756	GCTGAGGTCACCTGTGCGCCCAAGCCAGGAGGCGGACGCTCTCCCTCATTCCTCCCTA	815
Db	393	GCTGAGGTCACCTGTGCGCCCAAGCCAGGAGGCGGACGCTCTCCCTCATTCCTCCCTA	452
QY	816	CAGCCCAAGTGAACCCACCTGGGTCTCCAGCAACCCACCTCCACCGAGTCTTAAT	875
Db	453	CAGCCCAAGTGAACCCACCTGGGTCTCCAGCAACCCACCTCCACCGAGTCTTAAT	512
QY	876	GCCTTGGATCCATGAATTTGCCGCGGGGTGCACCCAGGGCTCAGAGAGAAGGTTTTGAG	935

Db 513 |||||GCCTGGATCATGAATTCGCGCGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGGAG 572
 QY 936 ACAGCTGCTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCACTCTCTTACCA 995
 Db 573 ACAGCTGCTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCACTCTTACCA 632
 QY 996 AGCTCTTCAAGCTCCCGACACATGAGTCTTCAAGCCAGGAGTCAACCAAGCCGCCAC 1055
 Db 633 AGCTCTTCAAGCTCCCGACACATGAGTCTTCAAGCCAGGAGTCAACCAAGCCGCCAC 692
 QY 1056 ATCCACAGAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGA 1115
 Db 693 ATCCACAGAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGA 752
 QY 1116 GCAGGAGGCTCTGGGACCAATGGCTCTGGATGAAACGTGAGAAATTTACTTACACAGA 1175
 Db 753 GCAGGAGGCTCTGGGACCAATGGCTCTGGATGAAACGTGAGAAATTTACTTACACAGA 812
 QY 1176 AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGACGCGTGGT 1235
 Db 813 AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCCCATGGGACGCGTGGT 872
 QY 1236 AGGAAAGCCGCCACAGCGGACACAGCCTACAGCCGCCACCCACCCATGGGAGCCTTC 1295
 Db 873 AGGAAAGCCGCCACAGCGGACACAGCCTACAGCCGCCACCCACCCATGGGAGCCTTC 932
 QY 1296 TGTGAGAGAGGCTCTGTTTCCATGAGCCCTCGGAAATAGAGATTTTAAACCAAAATTT 1355
 Db 933 TGTGAGAGAGGCTCTGTTTCCATGAGCCCTCGGAAATAGAGATTTTAAACCAAAATTT 992
 QY 1356 CACACAGCTGCTACTTCTACAAAGACCTCACCCAGAGCCAGATCCCTGTGCAAGAG 1415
 Db 993 CACACAGCTGCTACTTCTACAAAGACCTCACCCAGAGCCAGATCCCTGTGCAAGAG 1052
 QY 1416 AAGCTGGCTGATTTATGTGAGAGAGAAATCGAGGACATTTAATTTAGATCAGAGACTTAT 1475
 Db 1053 AAGCTGGCTGATTTATGTGAGAGAGAAATCGAGGACATTTAATTTAGATCAGAGACTTAT 1112
 QY 1476 TGGCCAGAGCTGGATACCAAGACCTCGCATAGTCACTACGAGGGGCTGCTGGAAT 1535
 Db 1113 TGGCCAGAGCTGGATACCAAGACCTCGCATAGTCACTACGAGGGGCTGCTGGAAT 1172
 QY 1536 TGGGAAGTCAACACTGCGCCAGGAGGTGAAGAGGCTGGGGAGAGGCGCAGCTGTATGG 1595
 Db 1173 TGGGAAGTCAACACTGCGCCAGGAGGTGAAGAGGCTGGGGAGAGGCGCAGCTGTATGG 1232
 QY 1596 GGAACCGCTTCCAGCATGCTTCTTCACTTCACTGAGAGAGCTGGCCCAAGTCCAAGTGGT 1655
 Db 1233 GGAACCGCTTCCAGCATGCTTCTTCACTTCACTGAGAGAGCTGGCCCAAGTCCAAGTGGT 1292
 QY 1656 GAGTCTCGCTGAGCTCATCGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1715
 Db 1293 GAGTCTCGCTGAGCTCATCGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1352
 QY 1716 CTTGCTTAGGCCAGAGCGGCTGCTTCTTCACTTCACTGAGTGGTGTAGATGAGCCAGGATGGT 1775
 Db 1353 CTTGCTTAGGCCAGAGCGGCTGCTTCTTCACTTCACTGAGTGGTGTAGATGAGCCAGGATGGT 1412
 QY 1776 CTTGAGGAGCCGAGTCTGAGCTCTGTGTGCACTGGAGCCACCCAGCCAGCCGCGGATGC 1835
 Db 1413 CTTGAGGAGCCGAGTCTGAGCTCTGTGTGCACTGGAGCCACCCAGCCAGCCGCGGATGC 1472
 QY 1836 ACTGCTGGGAGTCTGGGAGAAACTATACCTTCCGAGGATCCTTCTGATCAGGC 1895
 Db 1473 ACTGCTGGGAGTCTGGGAGAAACTATACCTTCCGAGGATCCTTCTGATCAGGC 1532
 QY 1896 TCGGACACAGCTCTGAGAACTCATTTCTCTTTGGAGCAGCAGCTGGGTAGAGT 1955
 Db 1533 TCGGACACAGCTCTGAGAACTCATTTCTCTTTGGAGCAGCAGCTGGGTAGAGT 1592
 QY 1956 CTTGGGCTTCTGAGTCCAGCAGGAGGAATTTCTCAGATATTTTCAAGATGAAAG 2015

Db 1593 CTTGGGCTTCTGAGTCCAGCAGGAGGAATATTTTACAGATATTTTACAGATGAAG 1652
 QY 2016 GCAAGCAATTAAGAGCTTTAGGTTGTTCAAAACAAAGAGCTCTGGGCCCTGTGTCT 2075
 Db 1653 GCAAGCAATTAAGAGCTTTAGGTTGTTCAAAACAAAGAGCTCTGGGCCCTGTGTCT 1712
 QY 2076 TGTGCTTGGGTGCTGCTGGCTGCACTTTGCTGATGAGCAGAGATGAAGCGGAGGA 2135
 Db 1713 TGTGCTTGGGTGCTGCTGGCTGCACTTTGCTGATGAGCAGAGATGAAGCGGAGGA 1772
 QY 2136 AAAATCTACACTGACTTCCAGACACCAACCTCTGTCTACATTAATCTTTGCCAGGC 2195
 Db 1773 AAAATCTACACTGACTTCCAGACACCAACCTCTGTCTACATTAATCTTTGCCAGGC 1832
 QY 2196 TCTCAAGCTCAGGCAATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGG 2255
 Db 1833 TCTCAAGCTCAGGCAATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGG 1892
 QY 2256 CATCTGGCAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGG 2315
 Db 1893 CATCTGGCAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGG 1952
 QY 2316 GGCCATCATCTCCACTTCTTGAAGATGGTATTTCTTCAAGAGCACCCCATCTCTGAG 2375
 Db 1953 GGCCATCATCTCCACTTCTTGAAGATGGTATTTCTTCAAGAGCACCCCATCTCTGAG 2012
 QY 2376 CTACAGCTTCATTCACCTCTGTTTCAAGAGTCTTTTGCAGCAATGTCTATGTCTTGA 2435
 Db 2013 CTACAGCTTCATTCACCTCTGTTTTCAGGAGTCTTTTGCAGCAATGTCTATGTCTTGA 2072
 QY 2436 GGATGAGAGGGGAGGAGGTAAACATTTTAAATTTGCAATCATAGATTTTGAAGAGCAGCTAGA 2495
 Db 2073 GGATGAGAGGGGAGGAGGTAAACATTTTAAATTTGCAATCATAGATTTTGAAGAGCAGCTAGA 2132
 QY 2496 AGCATATGGAATACATGGCTCTTTGGGCAATCAACACACCTTTTCTTATGGGCTGTT 2555
 Db 2133 AGCATATGGAATACATGGCTCTTTGGGCAATCAACACACCTTTTCTTATGGGCTGTT 2192
 QY 2556 AAGTATGAGGGGAGAGAGATGGAGAAATCTTTTCACTGCGGCTCTCTCAGGGAG 2615
 Db 2193 AAGTATGAGGGGAGAGAGATGGAGAAATCTTTTCACTGCGGCTCTCTCAGGGAG 2252
 QY 2616 GAACCTGATGCAAGTGGGCTCCCTGCTGAGCTGTGTGAGCCACACTCTCTGAGTGC 2675
 Db 2253 GAACCTGATGCAAGTGGGCTCCCTGCTGAGCTGTGTGAGCCACACTCTCTGAGTGC 2312
 QY 2676 CTTCACTGCTTGTACGAGACTCGGAAACAAACGTTCTTGAACAGTGTATGGCCCATTT 2735
 Db 2313 CTTCACTGCTTGTATGAGACTCGGAAACAAACGTTCTTGAACAGTGTATGGCCCATTT 2372
 QY 2736 CGAAGAAATGGGCATGTGTGTAGAAACACAGATGGAGCTCTTAGTGTGCACTTTCTGCAT 2795
 Db 2373 CGAAGAAATGGGCATGTGTGTAGAAACACAGATGGAGCTCTTAGTGTGCACTTTCTGCAT 2432
 QY 2796 TAAATTCAGCCGCCAGCTGGAAGAGCTTCAGCTGATTTAGGGCAGGACAGATCAAC 2855
 Db 2433 TAAATTCAGCCGCCAGCTGGAAGAGCTTCAGCTGATTTAGGGCAGGACAGATCAAC 2492
 QY 2856 ATGGAGCCGCCACATGAGTGGTCTTGTAGTGGGTCCAGTCAAGATGCTTATTTGGCA 2915
 Db 2493 ATGGAGCCGCCACATGAGTGGTCTTGTAGTGGGTCCAGTCAAGATGCTTATTTGGCA 2552
 QY 2916 GATTCCTTCTCCGCTCCTCAAGGTCAACAGAACTTGAAGAGCTTGAAGTGAAGTGAAG 2975
 Db 2553 GATTCCTTCTCCGCTCCTCAAGGTCAACAGAACTTGAAGAGCTTGAAGTGAAGTGAAG 2612
 QY 2976 CTCGCTGAGCCACTCTGCAAGTGAAGAGCTTGTGTAAAGCCCTGAGACGCCCTCGCTGCCT 3035
 Db 2613 CTCGCTGAGCCACTCTGCAAGTGAAGAGCTTGTGTAAAGCCCTGAGACGCCCTCGCTGCCT 2672
 QY 3036 CTTGAGAGCCCTGCGGCTGGCTGGCTCTCACAGCTGAGGACTGCAAGAGACCTTGC 3095
 Db 2673 CTTGAGAGCCCTGCGGCTGGCTGGCTCTCACAGCTGAGGACTGCAAGAGACCTTGC 2732

QY 3096 CTTTGGGCTGAGAGCCAAACAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGTCTAC 3155
 Db 2733 CTTTGGGCTGAGAGCCAAACAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGTCTAT 2792
 QY 3156 GGATGCTGAGAGCCAAACACCTTTGGCAGAGACTGAGACAGCGAGCTGCAAGCTACAGCG 3215
 Db 2793 GGATGCTGAGAGCCAAACACCTTTGGCAGAGACTGAGACAGCGAGCTGCAAGCTACAGCG 2852
 QY 3216 ACTGACGCTGCTGAGCTGTGGCTCAGCTGCTGACTGCTGCCAGGACCTGSCCTCTGTGCT 3275
 Db 2853 ACTGACGCTGCTGAGCTGTGGCTCAGCTGCTGACTGCTGCCAGGACCTGGCTCTGTGCT 2912
 QY 3276 TAGTGCCAGCCCGAGCCTGAAGGAGCTAGACCTGACAGCAAGAACCTTGGATGACCTTGG 3335
 Db 2913 TAGTGCCAGCCCGAGCCTGAAGGAGCTAGACCTGACAGCAAGAACCTTGGATGACCTTGG 2972
 QY 3336 CFTGCACTGCTCTGTGAGGGCTCAGGATCTGCTGCTGCAAACTCATACGCTGGGGCT 3395
 Db 2973 CFTGCACTGCTCTGTGAGGGCTCAGGATCTGCTGCTGCAAACTCATACGCTGGGGCT 3032
 QY 3396 GGACAGACAACTCTGAGTGTAGATGAGGAGGAGGAGCTGAGGGCCCTGGAGCAGAGAA 3455
 Db 3033 GGACAGACAACTCTGAGTGTAGATGAGGAGGAGGAGCTGAGGGCCCTGGAGCAGAGAA 3092
 QY 3456 ACCTCAGCTGCTCATCTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3515
 Db 3093 GCCTCAGCTGCTCATCTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3152
 QY 3516 GGATCGGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCGAGAGACTCGGATCAGA 3575
 Db 3153 GGATCGGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCGAGAGACTCGGATCAGA 3212
 QY 3576 GAGGGGGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGGAGCTGAGCAAGATCTT 3635
 Db 3213 GAGGGGGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGGAGCTGAGCAAGATCTT 3272
 QY 3636 CCCAATGCTGAGATTGACAGAGAAAGCTCCCGAGAGTAGTACCGGTGGAATCTTTGTG 3695
 Db 3273 CCCAATGCTGAGATTGACAGAGAAAGCTCCCGAGAGTAGTACCGGTGGAATCTTTGTG 3332
 QY 3696 CFTGCTTCTCTGCTCTCAAGGGACCTGATACGAAAGCTTTGGGACCTGACGATGA 3755
 Db 3333 CFTGCTTCTCTGCTCTCAAGGGACCTGATACGAAAGCTTTGGGACCTGACGATGA 3392
 QY 3756 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGGTAGTGAAGAAAGAACTTTGTA 3815
 Db 3393 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGGTAGTGAAGAAAGAACTTTGTA 3452
 QY 3816 CCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACCGGGTCTCTGCTTTGT 3875
 Db 3453 CCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACCGGGTCTCTGCTTTGT 3512
 QY 3876 GATGAGAGAGCGGTGACCGTTGAGATTGAAATTTGTGTGGACCACTTCTGGGTGA 3935
 Db 3513 GGTGAGAGAGCGGTGACCGTTGAGATTGAAATTTGTGTGGACCACTTCTGGGTGA 3572
 QY 3936 GATCAACCCACAGCAGCTGATGTGTGGAGGGCTCTGCTGGACATCAAGCTGTAGCC 3995
 Db 3573 GATCAACCCACAGCAGCTGATGTGTGGAGGGCTCTGCTGGACATCAAGCTGTAGCC 3632
 QY 3996 TGGAGCTGTGGAAGCTGTCACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCATGTGA 4055
 Db 3633 TGGAGCTGTGGAAGCTGTCACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCATGTGA 3692
 QY 4056 CACATCCCTGTTCCAAATGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAAGCCAGC 4115
 Db 3693 CACATCCCTGTTCCAAATGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAAGCCAGC 3752
 QY 4116 CAGGTGTGAGCTGCATCATACTGTCTGGAAACCCAGCTTCTCCCCCTTGGAGTCTCT 4175
 Db 3753 CAGGTGTGAGCTGCATCATACTGTCTGGAAACCCAGCTTCTCCCCCTTGGAGTCTCT 3812

QY 4176 CTTGAAAATGATTCATATGCCCTGCGCTTCATTCCCTCAGCTCTGTGGTGTGCTTTA 4235
 Db 3813 CTTGAAAATGATTCATATGCCCTGCGCTTCATTCCCTCAGCTCTGTGGTGTGCTTTA 3872
 QY 4236 CCACCGCTCCATCTCTGAGGAGTCACTTCCACTCTACCTGATCCCAAGTGAAGTCTC 4295
 Db 3873 CCACCGCTCCATCTCTGAGGAGTCACTTCCACTCTACCTGATCCCAAGTGAAGTCTC 3932
 QY 4296 CATTCGGAAGAACTGGAGCTCTGCTATCGAAGCCCTCGAAGAACCAAGCTGTTCTCGGA 4355
 Db 3933 CATTCGGAAGAACTGGAGCTCTGCTATCGAAGCCCTCGAAGAACCAAGCTGTTCTCGGA 3992
 QY 4356 GTTCTACGTTGGCCTCTGGGATCAGGATCAGGCTGCAAGTGAAGAACCAAGTGAAGTGA 4415
 Db 3993 GTTCTACGTTGGCCTCTGGGATCAGGATCAGGCTGCAAGTGAAGAACCAAGTGAAGTGA 4052
 QY 4416 GACTCTGTGTGGAGGCTTGGTGAACCAAGGAGATCTCATGCTGCAACTACTCTGAT 4475
 Db 4053 GACTCTGTGTGGAGGCTTGGTGAACCAAGGAGATCTCATGCTGCAACTACTCTGAT 4112
 QY 4476 CCCTCCAGCCCGCATAGCCGTACCTTTCACCTCTGGATGCCCGCAGTGTGCTGCATTTGT 4535
 Db 4113 CCCTCCAGCCCGCATAGCCGTACCTTTCACCTCTGGATGCCCGCAGTGTGCTGCATTTGT 4172
 QY 4536 GGACCACTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGTGTGCTTGGACAA 4595
 Db 4173 GGACCACTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGTGTGCTTGGACAA 4232
 QY 4596 ACTGATCGACAGTGTGAGCCAGGAGCAGTACGAGAGGGTGTGCTGCTGAGAACACAG 4655
 Db 4233 ACTGATCGACAGTGTGAGCCAGGAGCAGTACGAGAGGGTGTGCTGCTGAGAACACAG 4292
 QY 4656 GCCACGACAGTGGGAAGCTGTTTCAGCTTGAGCCAGTCTTGGGACCGGAAGTGAAGA 4715
 Db 4293 GCCACGACAGTGGGAAGCTGTTTCAGCTTGAGCCAGTCTTGGGACCGGAAGTGAAGA 4352
 QY 4716 TGGACTCTACCAAGCCCTGGAAGGAGACCCATCTCACCTCATTTATGGAACCTGCGGAGAA 4775
 Db 4353 TGGACTCTACCAAGCCCTGGAAGGAGACCCATCTCACCTCATTTATGGAACCTGCGGAGAA 4412
 QY 4776 GGGCAGAAAAAGGACTCTTGCCTCAGCAGCTGAAAGTATCAACACAGCCCTTGACC 4835
 Db 4413 GGGCAGAAAAAGGACTCTTGCCTCAGCAGCTGAAAGTATCAACACAGCCCTTGACC 4472
 QY 4836 CTTGAGTCTGCTGCTTGGCTGACCCCTCTTGGGTCTCAGTCTTCTCTGCAAAACAG 4895
 Db 4473 CTTGAGTCTGCTGCTTGGCTGACCCCTCTTGGGTCTCAGTCTTCTCTGCAAAACAG 4532
 QY 4896 TTGCCATCTGCTTGGCTTCCAGCACTAAAGTAAATGGAACCTTGTATGATGCTTGTGCTGG 4955
 Db 4533 TTGCCATCTGCTTGGCTTCCAGCACTAAAGTAAATGGAACCTTGTATGATGCTTGTGCTGG 4592
 QY 4956 GCATTTATGCTCATGCGGAGTGCACAGGGGGCCCGCTCAGGTGGCCCTAACAGCA 5015
 Db 4593 GCATTTATGCTCATGCGGAGTGCACAGGGGGCCCGCTCAGGTGGCCCTAACAGCA 4652
 QY 5016 TCTCAGGAAGTCTCATCTGAGCTGGCAAGACCCCTGACAGCTCATAGAGCTCATCT 5075
 Db 4653 TCTCAGGAAGTCTCATCTGAGCTGGCAAGACCCCTGACAGCTCATAGAGCTCATCT 4712
 QY 5076 GTTGCCACAGCAGCAAGCCCTAGAGCCCTCCGATCCATCCAGGGCCCAAGAGGAATA 5135
 Db 4713 GTTGCCACAGCAGCAAGCCCTAGAGCCCTCCGATCCATCCAGGGCCCAAGAGGAATA 4772
 QY 5136 GGAGGGAATGGAACCATTTGCTCTGCTGTGTCAAGGGTGGCCCCCAAAATTTGGGGT 5195
 Db 4773 GGAGGGAATGGAACCATTTGCTCTGCTGTGTCAAGGGTGGCCCCCAAAATTTGGGGT 4832
 QY 5196 TCAGGTGGAGGGCCAGCTGGAATTTCTGGCTTGTGTACAGGAAGATCTCAAGAGCAAGCC 5255
 Db 4833 TCAGGTGGAGGGCCAGCTGGAATTTCTGGCTTGTGTACAGGAAGATCTCAAGAGCAAGCC 4892
 QY 5256 AACAGGTAAAGTGAAGGAAGTTTATTTCAGAAAAATAAGAGGATATCAGCTCTTTTAG 5315

Db 4893 AACACAGTAAAGTGAAGAAAGTTTATTCAGAAAAATAAAGAGTATCACTGCTCTTTTAG 4952
Qy 5316 AATTTGCTAGCAGGCTTCCAGTCTTTTACACAGAAAAACCCCTATAAAATTAATAATTTT 5375
Db 4953 AATTTGCTAGCAGACTTCCAGTCTTTTACACAGAAAAACCCCTATAAAATTAATAATTTT 5012
Qy 5376 ACTTAAATTAAGAAATTAATAAAATTAATAAAATTAATAAAATTAATAAAATTAATAAA 5435
Db 5013 ACTTAAATTAAGAAATTAATAAAATTAATAAAATTAATAAAATTAATAAAATTAATAAA 5072
Qy 5436 GTTA 5439
Db 5073 GTTA 5076

RESULT 5
AX207661 5122 bp DNA linear PAT 31-AUG-2001
LOCUS
DEFINITION Sequence 29 from Patent WO0157085.
ACCESSION AX207661
VERSION AX207661.1 GI:15422343
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baughn, M.R., Au-Young, J. and Yue, H.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0157085-A 29 09-AUG-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..5122
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 4250893CB1"
BASE COUNT 1184 a 1460 c 1401 g 1077 t
ORIGIN

Query Match 87.5%; Score 4765.6; DB 6; Length 5122;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 4933; Conservative 0; Mismatches 24; Indels 133; Gaps 2;

Qy 35 GAGCAGCAGCCGGGGCTCCACTCTGGGTTCTGAAGCCCATTCCTCTCTGGGGCTC 94
Db 1 GAGCCAGCAGCCGGGGCTCCACTCTGGGTTCTGAAGCCCATTCCTCTCTGGGGCTC 60

Qy 95 CTCACACCCACCTCTTCTCAGCCTTGCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACC 154
Db 61 CTCACACCCACCTCTTCTCAGCCTTGCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACC 120

Qy 155 CAGGAGAGGAAGAATCTGAGGAACAACAGAACAGTGAGGCTTGCCCAACCCCATCTCCC 214
Db 121 CAGGAGAGGAAGAATCTGAGGAACAACAGAACAGTGAGGCTTGCCCAACCCCATCTCCC 180

Qy 215 GTACACATCTCCCTCACCCTCACCCTCCCTGCTGAGCCCTGAGCCCATCCAGGAC 274
Db 181 GTACACATCTCCCTCAGTAAACACCTCCCTGCTGAGCCCTGAGCCCATCCAGGAC 240

Qy 275 CTCCTTATCAGCTGACTTCTTCAGTGTCTTGAGGCCCCCTCTGGGCTCTCTCCCTCCCCT 334
Db 241 CTCCTTATCAGCTGACTTCTTCAGTGTCTTGAGGCCCCCTCTGGGCTCTCTCCCTCCCCT 300

Qy 335 GGCTTTTCTACACATCCCTCTATCGGCTCTATCTGAGTGCCCTGGGATTTATAA 394
Db 301 GGCTTTTCTACACATCCCTCTATCGGCTCTATCTGAGTGCCCTGGGATTTATAA 360

Qy 395 AACTGGGTTCCGATCTGATGAAGACGGTGAAGCCNAGCAAGGACAGCAGCTGTT 454
Db 361 AACTGGGTTCCGATCTGATGAAGACGGTGAAGCCNAGCAAGGACAGCAGCTGTT 420

Qy 455 CTCTGCTGCTGATACCTTACACCTTGGGAACATCCCCAGACACCTCTTAACCTCG 514
Db 421 CTCTGCTGCTGATACCTTACACCTTGGGAACATCCCCAGACACCTCTTAACCTCG 480

Qy 515 GGACAGAGATGGCTGGCGAGCCTGGGGCCGCTGGCTGTGTACTTTGAGTTCTCTGAAGA 574
Db 481 GGACAGAGATGGCTGGCGAGCCTGGGGCCGCTGGCTGTGTACTTTGAGTTCTCTGAAGA 540

Qy 575 AGGAGGAGCTGAAGAGGTTCCAGCTTCTGCTGCGCAATAAAGCGCACTCCAGGAGCTTT 634
Db 541 AGGAGGAGCTGAAGAGGTTCCAGCTTCTGCTGCGCAATAAAGCGCACTCCAGGAGCTTT 600

Qy 635 CGGCTGAGACACCCCGCTCAGCCAGAGAAAGAGTGGCATGGAGTGGCTCTGTACCTGG 694
Db 601 CGGCTGAGACACCCCGCTCAGCCAGAGAAAGAGTGGCATGGAGTGGCTCTGTACCTGG 660

Qy 695 TGGCTCAGTATGGGAGCAGGGGCTGGGACCTAGCCCTTCCATACCTGGGAGCAGATGG 754
Db 661 TGGCTCAGTATGGGAGCAGGGGCTGGGACCTAGCCCTTCCATACCTGGGAGCAGATGG 720

Qy 755 GGCTGAGGTCACTGTGCGCCCAAGCCAGGAAGGGGCGAGGCCACTCTCTCCCTCATTTCCCT 814
Db 721 GGCTGAGGTCACTGTGCGCCCAAGCCAGGAAGGGGCGAGGCCACTCTCTCCCTCATTTCCCT 780

Qy 815 ACAGCCCAAGTGAACCCCGCTGGGCTTCCAGAGCAACCCACCTCCACCCGAGTGTCTAA 874
Db 781 ACAGCCCAAGTGAACCCCGCTGGGCTTCCAGAGCAACCCACCTCCACCCGAGTGTCTAA 840

Qy 875 TGCCCTGGATCCATGAATTCGCGGGGTGCACCCAGGCTCAGAGAGAGGGTTTCA 934
Db 841 TGCCCTGGATCCATGAATTCGCGGGGTGCACCCAGGCTCAGAGAGAGGGTTTCA 900

Qy 935 GACAGCTGCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCCTCATCTCTTACC 994
Db 901 GACAGCTGCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCCTCATCTCTACC 960

Qy 995 AAGCTCTTCCAAAGTCCCGCAGACATAGTCTTCCAGCCAGAGAGTACCCACAGCCCCCA 1054
Db 961 AAGCTCTTCCAAAGTCCCGCAGACATAGTCTTCCAGCCAGAGAGTACCCACAGCCCCCA 1020

Qy 1055 CATCCACAGCAGTGTGGGGAGCTGGGATCCACCTCAGCCAGCTAGCACCCAGAG 1114
Db 1021 CATCCACAGCAGTGTGGGGAGCTGGGATCCACCTCAGCCAGCTAGCACCCAGAG 1080

Qy 1115 AGCAGAGGCTCCTGGGACCCCAATGGCTCTCGATGAACAGTCAGGAATTTACTACACAG 1174
Db 1081 AGCAGAGGCTCCTGGGACCCCAATGGCTCTCGATGAACAGTCAGGAATTTACTACACAG 1140

Qy 1175 AAATCAGAGAAAGAGAGAGAGAAATCAGAAAGGCGAGGCCCTCATGGGAGCGGTGG 1234
Db 1141 AAATCAGAGAAAGAGAGAGAGAAATCAGAAAGGCGAGGCCCTCATGGGAGCGGTGG 1200

Qy 1235 TAGGAACCCCCCAGGGGACACAGCCTACAGCCACCCACACCCATCTGGAGCCTT 1294
Db 1201 TAGGAACCCCCCAGGGGACACAGCCTACAGCCACCCACACCCATCTGGAGCCTT 1260

Qy 1295 CTGTGAGAGAGAGCTCTGTTCACATGGCCCTGGAAAAATCAGGATTTTAAACCAAAAT 1354
Db 1261 CTGTGAGAGAGAGCTCTGTTCACATGGCCCTGGAAAAATCAGGATTTTAAACCAAAAT 1320

Qy 1355 TCACACAGCTGCTATTTCTACAAAGACCTCAACCCAGAGAGCAAGATCCCTGGTCAAGA 1414
Db 1321 TCACACAGCTGCTATTTCTACAAAGACCTCAACCCAGAGAGCAAGATCCCTGGTCAAGA 1380

Qy 1415 GAAGCTGGCTGATTTATGTGGAGAGAAATCAGGACATTTAATTTGAGATCAGAGCTTAT 1474
Db 1381 GAAGCTGGCTGATTTATGTGGAGAGAAATCAGGACATTTAATTTGAGATCAGAGCTTAT 1440

Qy 1475 TTGGCCCGGCTGGATACCCCAAGAACCTCGCATAGTCTATCTGACAGGGGCTGTGGAA 1534
Db 1441 TTGGCCCGGCTGGATACCCCAAGAACCTCGCATAGTCTATCTGACAGGGGCTGTGGAA 1500

Qy 1535 TTGGGAAGTCAACACTGGCCAGGAGGTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATG 1594

QY 3755 ACTTCTGGGGCCCCACGGGGCCCTGTGGCTACTGAGGTAGTGTGACAAAGAAAAGAACTTCT 3814
Db 3721 ACTTCTGGGGCCCCACGGGGCCCTGTGGCTACTGAGGTAGTGTGACAAAGAAAAGAACTTCT 3780
QY 3815 ACCGAGTTCACTTCCCTGTAGTGGCTCCTACCGCTGGGCCCAACACGGGTCTCTGCTTGG 3874
Db 3781 ACCGAGTTCACTTCCCTGTAGTGGCTCCTACCGCTGGGCCCAACACGGGTCTCTGCTTGG 3840
QY 3875 TGATGAGAGAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTG 3934
Db 3841 TGNATGAGAGAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTG 3900
QY 3935 AGATCAACCCACACGACAGCTGGATGGTGGCAGGGCCTCTGTCTGGACATCAAGGCTGAGC 3994
Db 3901 AGATCAACCCACACGACAGCTGGATGGTGGCAGGGCCTCTGTCTGGACATCAAGGCTGAGC 3960
QY 3995 CTGGAGCTGTGGNAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGG 4054
Db 3961 CTGGAGCTGTGGNAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGG 4020
QY 4055 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCGGAGAACCCAG 4114
Db 4021 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCGGAGAACCCAG 4080
QY 4115 CCAGGGTGGAGCTGCATCATAGTTCTGGAAAACCCCACTTCTCCCTCTGGGAGTCC 4174
Db 4081 CCAGGGTGGAGCTGCATCATAGTTCTGGAAAACCCCACTTCTCCCTCTGGGAGTCC 4140
QY 4175 TCCTGAAATGATCCATATATGCCCTGGCTTCATTCCTCGTCACTCTGTGGTGTGCTTT 4234
Db 4141 TCCTGAAATGATCCATATATGCCCTGGCTTCATTCCTCGTCACTCTGTGGTGTGCTTT 4200
QY 4235 ACCACGGCTGCATCTGAGGAAGTCACTTCCACCTCACTGATCCCAAGTCACTGCT 4294
Db 4201 ACCACGGCTGCATCTGAGGAAGTCACTTCCACCTCACTGATCCCAAGTCACTGCT 4260
QY 4295 CCATTCGG----- 4302
Db 4261 CCATTCGGAAGGCCATAGATGATCTAGAAATGAAATTCAGTTTGTGCGAATCCACAAGC 4320
QY 4303 ----- 4302
Db 4321 CACCCCGCTGACCCCACTTATATGGGCTGTCTTACATGCTGTGCTGGTTCAG 4380
QY 4303 -----AAGGAAGTGGAGCTGCTATCGAAGCCCTGGAGAGCC 4342
Db 4381 GGATGTGGAATACTCCCAAGGAAGTGGAGCTCTGCTATCGAAGCCCTGGAGAGAGCC 4440
QY 4343 AGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAG 4402
Db 4441 AGCTGTTCTCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAG 4500
QY 4403 ACAAGAAAGATGAGACTCTGGTGTGGAGGCTTGTGTGAACACAGAGATCTCATCCCTG 4462
Db 4501 ACAAGAAAGATGAGACTCTGGTGTGGAGGCTTGTGTGAACACAGAGATCTCATCCCTG 4560
QY 4463 CAATCTACTGATCCCTCAGCCCGCATAGCCGCTACCTTCACTCTGGATGCCCGCAGT 4522
Db 4561 CAATCTACTGATCCCTCAGCCCGCATAGCCGCTACCTTCACTCTGGATGCCCGCAGT 4620
QY 4523 TGCTGCATTTTGGGACCATGATCGAGAGCAGTGTATAGCCCGAGTGACATCGTGGAGG 4582
Db 4621 TGCTGCATTTTGGGACCATGATCGAGAGCAGTGTATAGCCCGAGTGACATCGTGGAGG 4680
QY 4583 TTGCTTTGGACAAACTGCATGGAAGTGTGTGAGCCAGGAGCAGTACGAGAGGCTCTGG 4642
Db 4681 TTGCTTTGGACAAACTGCATGGAAGTGTGTGAGCCAGGAGCAGTACGAGAGGCTCTGG 4740
QY 4643 CTGAGAACACGAGGCCAGCCAGATCGGGAAGTGTTCAGTTGAGCCAGTCTCTGGAGC 4702
Db 4741 CTGAGAACACGAGGCCAGCCAGATCGGGAAGTGTTCAGTTGAGCCAGTCTCTGGAGC 4800

QY 4703 GGAAGTGCAAGATGGACTCTACCAAGCCCTGAAGAGAGACCCATCCTACCTCATATTGG 4762
Db 4801 GGAAGTGCAAGATGGACTCTACCAAGCCCTGAAGAGAGACCCATCCTACCTCATATTGG 4860
QY 4763 AACTCTGGGAGAGAGGAGCAAAAGAGGACTCTCCCACTCAGCAGCTGAAAGTATCAACA 4822
Db 4861 AACTCTGGGAGAGAGGAGCAAAAGAGGACTCTCCCACTCAGCAGCTGAAAGTATCAACA 4920
QY 4823 CCAGCCCTTGACCTTGAGTCTCTGGCTTGGCTGACCCCTTCTTTGGGTCTCAGTTCTTT 4882
Db 4921 CCAGCCCTTGACCTTGAGTCTCTGGCTTGGCTGACCCCTTCTTTGGGTCTCAGTTCTTT 4980
QY 4883 CTCTGCAAAACAAAGTTGCCATCTGGTTTGGCTTCCAGCATAAAGTAATGGAACCTTTGATG 4942
Db 4981 CTCTGCAAAACAAAGTTGCCATCTGGTTTGGCTTCCAGCATAAAGTAATGGAACCTTTGATG 5040
QY 4943 ATGCTCTTGGCTGGGCAATATGCTCATGCCAGGATGCCACAGGGGGCC 4992
Db 5041 GATGCTTGGCTGGGCAATATGCTCCA-GCCAGGATGCCAGGGGGGCC 5089

RESULT 6
AF229059 Homo sapiens NAC-alpha splice variant (NAC) mRNA, complete cds, alternatively spliced.
LOCUS AF229059
DEFINITION AF229059.1 GI:12656104
ACCESSION
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE: 1 (bases 1 to 4290)
AUTHORS Chu, Z.-L., Pio, F., Xie, Z., Godzik, A. and Reed, J.C.
TITLE NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c pathway for apoptosis
JOURNAL Unpublished
RENEAL 2 (bases 1 to 4290)
AUTHORS Chu, Z.-L. and Reed, J.C.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
Location/Qualifiers
1..4290
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/cell_type="activated Jurkat T cells"
1..4290
/gene="NAC"
1..4290
/gene="NAC"
/note="Apaf-1/Ced-4 family apoptosis regulator; contains NB-ARC domain, CARD domain, and leucine-rich repeats (LRR)"
/codon_start=1
/product="NAC-alpha splice variant"
/protein_id="AAK00748.1"
/db_xref="GI:12656105"
translation="MAGAWGLACVLFKKEELKEFOLLANKAHSSSSGTEPAQ PEKTSMEVSYLVAYQYEQRAWDLALHTWEQGLRSICAQAGAGHSPPSPFSE PHLGSPQSTAVLPMIHELPACTGSSERRVLRQPDTSGRWRREISAGLLYQAL PPSPPHSPSPSNAPTSTAVLGSWGPQPSPSLAPREQEAPFTQWLPDETSGIYYTE IRERERKSEKRPAAVAVGTPQATSLQPHHPSPVRESCLSTWPKNEPQNFQ KFTOLLQRPHPRSQDPLVKRSPDYVEENRHLIEIRDLFGPLDTPQEPVILQV AAGICKSLARQKEAWGEGYGRFOHVEFVESCRLAQSKVSLAEIIGKGTATP APIRQLRPERLLFILGVDEPFWLVEPSELCLHNSQOPADALGSLGKILP EASFLITATLQNLIPSEQRWVEVLESSESKRYFYFRFDEQAIRAFRLVK SNKELWALCLVFWSPQCLMQQMKRKLTLTKTTLCLLYLAQALQALQPLG QLRDLCSAAEGIWQKTLFSPDDLRLKHGLDGLAIISTFLKMGILQHPILSYSL

gene
CDS

CFQEPFAAMSYVLEDEKGRKHSNCIDLEKTLWAYIGHLFGCASTRFLGLLSDBG
 EREMENIFHICRISQSNLMQWPSLOLLIQLPHSLHCLYETRNKFTLQWMAHFE
 MCMCVETIDELLAVCTFCIKFSRVKKLIQI EGROHRSWSPWVWVPTDAYWO
 ILPSVLKVTNRLKELDLGNSLSHSAVSLCKTLRPRCLLETFLRAGCGLTAECDK
 LAFGLRANOTLTELDSFNVLDAGNKHUCQRLRQPSCKLRLQLVSCGLSDCCQDL
 ASVLSASPSKELDLQONLDDVGVRLCEGLRHPACKLIRGLQDOTLISDEMROELR
 ALIEQKXPQLLIIFSRKPSVMTTEGLDGTGEMNSLTKRQLGSRASHVAQANLK
 LLDVSKIFPIABIAEBSSEVVELLCVPSASQDLHTKPLGTDDDDFWPGTGPVAT
 EVDVDEKNLYRVHPFVAGSVRPWNTGLCFVMEAVTEVEFCVWDOFLGEIINPOHSM
 VAGPLDDIKAEPGAEAHLPHPFVALOGSHVDTSLFQMAHFEKGLLEKPARVELH
 IVLENPSFPLGVLKHNALRFPVTSVLDLHRHPEEVTFHYLILIPSCSIRKE
 LELCYRSPGEDOLFSEFYVGHLSGIRLQDKDBETLWMEALVXPGLDMPATLIPP
 ARIAVPSDQQLHFLVQYREQLTARTVSEVLDLKLHGLVQVLSQEQYERVLAEATR
 PSQMRKLFSLSSWDRCKDKGLYOALKETHPHLIMELWEKSKGLPLSS"

BASE COUNT 1011 a 1189 c 1209 g 881 t

ORIGIN

Query Match 78.8%; Score 4290; DB 9; Length 4290;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	523	ATGGCTGGCGGAGCCTGGGGCGCCTGGCTGTACTTGGAGTTCCTGAAGAGGAG	582
DB	1	ATGGCTGGCGGAGCCTGGGGCGCCTGGCTGTACTTGGAGTTCCTGAAGAGGAG	60
QY	583	CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	642
DB	61	CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	120
QY	643	ACACCGCTCAGCAGAGAAGCAGTGTGAGGTGGCTCGTACCTGTGGTCTCAG	702
DB	121	ACACCGCTCAGCAGAGAAGCAGTGTGAGGTGGCTCGTACCTGTGGTCTCAG	180
QY	703	TATGGGAGCAGCGGCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG	762
DB	181	TATGGGAGCAGCGGCTGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG	240
QY	763	TCACTGTGCGCCCAAGCCAGGAAGGGCAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	822
DB	241	TCACTGTGCGCCCAAGCCAGGAAGGGCAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	300
QY	823	AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACCTCCACGCGAGTGTAAATGCCCTGG	882
DB	301	AGTGAACCCCACTGGGGTCTCCAGCCAAACCCACCTCCACGCGAGTGTAAATGCCCTGG	360
QY	883	ATCCATGAATTGGCGGGGGTGCAACCCAGGGCTCAGAGAAGGGTTTGTAGACAGCTG	942
DB	361	ATCCATGAATTGGCGGGGGTGCAACCCAGGGCTCAGAGAAGGGTTTGTAGACAGCTG	420
QY	943	CCTGCACATCTGGACGCGCTGGAGAAATCTCTGCTCTACTCTCTACCAAGCTCTT	1002
DB	421	CCTGCACATCTGGACGCGCTGGAGAAATCTCTGCTCTACTCTCTACCAAGCTCTT	480
QY	1003	CCAAGTCTCCAGACCATGAGTCTCCAAGCAGGAGTCAACCAAGCCGCCCATCCACA	1062
DB	481	CCAAGTCTCCAGACCATGAGTCTCCAAGCAGGAGTCAACCAAGCCGCCCATCCACA	540
QY	1063	GCAGTCTGGGAGCTGGGGATCCCCACCTCAGCCAGCCTAGCACCCAGAGAGGAG	1122
DB	541	GCAGTCTGGGAGCTGGGGATCCCCACCTCAGCCAGCCTAGCACCCAGAGAGGAG	600
QY	1123	GCTCTGGGACCAATGGCTCTGGATGAAACGTACAGGAATTTACTACACGAATTCAGA	1182
DB	601	GCTCTGGGACCAATGGCTCTGGATGAAACGTACAGGAATTTACTACACGAATTCAGA	660
QY	1183	GAAAGAGAGAGAGAGAAATCAGAGAAGGCAGGCCCCCATGGGACGGTGGTAGGAACG	1242
DB	661	GAAAGAGAGAGAGAGAAATCAGAGAAGGCAGGCCCCCATGGGACGGTGGTAGGAACG	720
QY	1243	CCCCCAGCGGCACACCGAGCTTACAGCCGCCACACCCATGGAGGCTTCTGTGAGA	1302
DB	721	CCCCCAGCGGCACACCGAGCTTACAGCCGCCACACCCATGGAGGCTTCTGTGAGA	780

QY	1303	GAGAGCCTCTGTTCCACATGCGCCTGGAAAAATAGGATTTTAAACAAAAATTACACAG	1362
DB	781	GAGAGCCTCTGTTCCACATGCGCCTGGAAAAATAGGATTTTAAACAAAAATTACACAG	840
QY	1363	CTGCTACTTCTTACAAGACCTTCAACCCAGAGACCCAGATCCCTGGTCAAGAGAAGCTGG	1422
DB	841	CTGCTACTTCTTACAAGACCTTCAACCCAGAGACCCAGATCCCTGGTCAAGAGAAGCTGG	900
QY	1423	CCTGATTTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA	1482
DB	901	CCTGATTTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA	960
QY	1483	GSCCTGGATACCCAAAGACCTCGCATAGTCATATGCAAGGGGCTGTGGAAATTTGGGAAG	1542
DB	961	GSCCTGGATACCCAAAGACCTCGCATAGTCATATGCAAGGGGCTGTGGAAATTTGGGAAG	1020
QY	1543	TCAACACTGGCGAGCGGTGAAGAAAGCCTGGGGAGAGCGCAGCTGTATTTGGGACCGC	1602
DB	1021	TCAACACTGGCGAGCGGTGAAGAAAGCCTGGGGAGAGCGCAGCTGTATTTGGGACCGC	1080
QY	1603	TTCCAGCATGTCTTCTACTTTCAGCTGTCAGAGAGCTGGCCCAAGTCCCAAGGTGGTGAAGTCTC	1662
DB	1081	TTCCAGCATGTCTTCTACTTTCAGCTGTCAGAGAGCTGGCCCAAGTCCCAAGGTGGTGAAGTCTC	1140
QY	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1722
DB	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1200
QY	1723	AGCCAGAGCGGCTGCTCTTCTCTCGATGGTGTAGATGAGCCAGGATGGTCTTGGCAG	1782
DB	1201	AGCCAGAGCGGCTGCTCTTCTCTCGATGGTGTAGATGAGCCAGGATGGTCTTGGCAG	1260
QY	1783	GAGCGGAGTTCGAGCTCTGTCTGCACTGGAGCCAGCCACAGCGCGGAGTGCATGCTGTG	1842
DB	1261	GAGCGGAGTTCGAGCTCTGTCTGCACTGGAGCCAGCCACAGCGCGGAGTGCATGCTGTG	1320
QY	1843	GGCAGTTTGTGGGAAAAATATATCTCCGAGGCGATCTTCTTGATCAGCGCTCGGACC	1902
DB	1321	GGCAGTTTGTGGGAAAAATATATCTCCGAGGCGATCTTCTTGATCAGCGCTCGGACC	1380
QY	1903	ACAGCTCTGCAAGACCTCATCTTCTTGGAGCAGGCGACGTTGGGTAGAGGTCCTGGGG	1962
DB	1381	ACAGCTCTGCAAGACCTCATCTTCTTGGAGCAGGCGACGTTGGGTAGAGGTCCTGGGG	1440
QY	1963	TTCTCTGAGTCAGCAGGAAGGAATTTCTACAGATATTTTACAGATGAAAGGCAAGCA	2022
DB	1441	TTCTCTGAGTCAGCAGGAAGGAATTTCTACAGATATTTTACAGATGAAAGGCAAGCA	1500
QY	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCCC	2082
DB	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCCC	1560
QY	2083	TGGGTGCTGGGCTGCACTTGTGCTGATGAGCAGATGAAGCGGAAGAAAAAATC	2142
DB	1561	TGGGTGCTGGGCTGCACTTGTGCTGATGAGCAGATGAAGCGGAAGAAAAAATC	1620
QY	2143	ACACTGACTTCCAGACACCAACACCTCTGTCTTACATTTACCTTGGCCAGGCTCTCCAA	2202
DB	1621	ACACTGACTTCCAGACACCAACACCTCTGTCTTACATTTACCTTGGCCAGGCTCTCCAA	1680
QY	2203	GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGGTGTGTGGGCAATCTGG	2262
DB	1681	GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGGTGTGTGGGCAATCTGG	1740
QY	2263	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC	2322
DB	1741	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC	1800
QY	2323	ATCTCCACTTCTTGAAGATGGTATTTCTTCAAGAGCAGCCCATCTCTGAGCTTACAGC	2382
DB	1801	ATCTCCACTTCTTGAAGATGGTATTTCTTCAAGAGCAGCCCATCTCTGAGCTTACAGC	1860
QY	2383	TTCAATTACCTCTGTTTCCAAAGAGTTCCTTTTGGCAGCAATGTCTATGTCTTGGAGGATGAG	2442

1861	Db	 TTCAATTCACTCTGTTTTTCCAAAGAGTCTTTTGGCAGCAATGTCCTATGCTTTGGAGGATGAG	1920
2443	Qy	AAGGGGAGAGGTAAACATCTTAATTGCATCATAGATTTTGGAAGAAGCGCTAGAAGCANAT	2502
1921	Db	AAGGGGAGAGGTAAACATCTTAATTGCATCATAGATTTTGGAAGAAGCGCTAGAAGCANAT	1980
2503	Qy	GGAATACATGGCCCTGTTTGGGGCATCAACACACGTTTCCCTATTGGGCCCTGTTTAAGTGAT	2562
1981	Db	GGAATACATGGCCCTGTTTGGGGCATCAACACACGTTTCCCTATTGGGCCCTGTTTAAGTGAT	2040
2563	Qy	GAGGGGAGAGAGATGGGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG	2622
2041	Db	GAGGGGAGAGAGATGGGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG	2100
2623	Qy	ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGTGCAGCCACACTCTCTGAGAGTCCCTCCAC	2682
2101	Db	ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGTGCAGCCACACTCTCTGAGAGTCCCTCCAC	2160
2683	Qy	TGCTTGTACGAGACTCGGAAACAAACAGTTTCTTGACACAAGTGATGGCCCCATTTCGAAGAA	2742
2161	Db	TGCTTGTACGAGACTCGGAAACAAACAGTTTCTTGACACAAGTGATGGCCCCATTTCGAAGAA	2220
2743	Qy	ATGGGCATGTGTGTAGAAAAACAGACATGGAGCTCTTAGTGTGCATTTTCTGCATTAATTC	2802
2221	Db	ATGGGCATGTGTGTAGAAAAACAGACATGGAGCTCTTAGTGTGCATTTTCTGCATTAATTC	2280
2803	Qy	AGCCGCCACGTGAAGAAGCTTCACGTCATTGAGGGCAGGCAGCACAGATCAACATGTGAGC	2862
2281	Db	AGCCGCCACGTGAAGAAGCTTCACGTCATTGAGGGCAGGCAGCACAGATCAACATGTGAGC	2340
2863	Qy	CCCACCATGTGTAGTCTCTGTTTCAGGTGGGTCCCAGTCCACAGATGCCTATTGGCAGATTCCTC	2922
2341	Db	CCCACCATGTGTAGTCTCTGTTTCAGGTGGGTCCCAGTCCACAGATGCCTATTGGCAGATTCCTC	2400
2923	Qy	TTCTCCGTCTCAAGGTCAACAGAAAACTGTGAAGAGCTGGACCTTAATGTGAAAACTCGCTG	2982
2401	Db	TTCTCCGTCTCAAGGTCAACAGAAAACTGTGAAGAGCTGGACCTTAATGTGAAAACTCGCTG	2460
2983	Qy	AGCCACCTCTCAGTGAAGAGTCTTTGTGAAGACCTGTAGAGACCCCTGCCTGCCTCTCGGAG	3042
2461	Db	AGCCACCTCTCAGTGAAGAGTCTTTGTGAAGACCTGTAGAGACCCCTGCCTGCCTCTCGGAG	2520
3043	Qy	ACCTCGGGTTGGCTGTGGGCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG	3102
2521	Db	ACCTCGGGTTGGCTGTGGGCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG	2580
3103	Qy	CTGAGAGCCAAACAGACCCTGACCGAGGTGACCTGTAGGCTTCAATGTGCTCACGATGCT	3162
2581	Db	CTGAGAGCCAAACAGACCCTGACCGAGGTGACCTGTAGGCTTCAATGTGCTCACGATGCT	2640
3163	Qy	GGAGCCAAACACTTTGCCAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTCGAC	3222
2641	Db	GGAGCCAAACACTTTGCCAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTCGAC	2700
3223	Qy	CTGTCTCAGCTGTGGCCCTCACGCTCTGACTGTCTGCCAGGACCTGGGCTCTGTGCTTAGTGCC	3282
2701	Db	CTGTCTCAGCTGTGGCCCTCACGCTCTGACTGTCTGCCAGGACCTGGGCTCTGTGCTTAGTGCC	2760
3283	Qy	AGCCCCAGCCTGAAGAGGACTAGACCTGTGACGAGAAACACTTGGATGAGTTGGCGTGCGA	3342
2761	Db	AGCCCCAGCCTGAAGAGGACTAGACCTGTGACGAGAAACACTTGGATGAGTTGGCGTGCGA	2820
3343	Qy	CTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAG	3402
2821	Db	CTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAG	2880
3403	Qy	ACAACTCTGAGTGATGAGTGAAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG	3462
2881	Db	ACAACTCTGAGTGATGAGTGAAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG	2940
3463	Qy	CTGCTCATCTTTCAGCAGACGGAAACCAAGTGTGTATGATCCCCCTACTGAGGGCTGGATACG	3522

2941	Db	CTGCTCATCTTCACGACAGCGGAAACCAAGTGTGATGACCCCTTACTGAGGCGCTGGATACG	3000
3523	Qy	GGAGAGATGATTAATAGCACATCCTCACTCAAGCGCGACAGACTCGGATCAGAGAGGGCG	3582
3001	Db	GGAGAGATGATTAATAGCACATCCTCACTCAAGCGCGACAGACTCGGATCAGAGAGGGCG	3060
3583	Qy	GCTTCCCATGTTGCTCAGGCTAATCTCAAACCTCTCGAAGTGAAGACTCTTCCCAATT	3642
3061	Db	GCTTCCCATGTTGCTCAGGCTAATCTCAAACCTCTCGAAGTGAAGACTCTTCCCAATT	3120
3643	Qy	GCTGAGATTGCAGAGAAAGCTCCCCAGAGGTAGTACCGTGGAACTCTTGTGCGTGCCT	3702
3121	Db	GCTGAGATTGCAGAGAAAGCTCCCCAGAGGTAGTACCGTGGAACTCTTGTGCGTGCCT	3180
3703	Qy	TCCTCTGCCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGGACTGACGATGACTTCTGG	3762
3181	Db	TCCTCTGCCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGGACTGACGATGACTTCTGG	3240
3763	Qy	GGCCCCACGGGCGCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTTGTACCGAGTT	3822
3241	Db	GGCCCCACGGGCGCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTTGTACCGAGTT	3300
3823	Qy	CACCTCCCTGTAGCTGCTCCCTACCGCTGCGCCAAACACCGGTCTCTGCTTTGTGATGAGA	3882
3301	Db	CACCTCCCTGTAGCTGCTCCCTACCGCTGCGCCAAACACCGGTCTCTGCTTTGTGATGAGA	3360
3883	Qy	GAAGCGGTGACCTTGAGATTGAATTCGTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC	3942
3361	Db	GAAGCGGTGACCTTGAGATTGAATTCGTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC	3420
3943	Qy	CCACACACAGCTGGATGTGGCAGGCGCTCTCTCTGGACATCAAGGCTGAGCCTGGAGCT	4002
3421	Db	CCACACACAGCTGGATGTGGCAGGCGCTCTCTCTGGACATCAAGGCTGAGCCTGGAGCT	3480
4003	Qy	GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAGGGGGCCATGTGGACACATCC	4062
3481	Db	GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAGGGGGCCATGTGGACACATCC	3540
4063	Qy	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGCCAGGGTG	4122
3541	Db	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGCCAGGGTG	3600
4123	Qy	GAGCTGCATCACAATAGTTCTGGAACCCACAGCTTCTCCCTTTGGAGTCTCTCGTGA	4182
3601	Db	GAGCTGCATCACAATAGTTCTGGAACCCACAGCTTCTCCCTTTGGAGTCTCTCGTGA	3660
4183	Qy	ATGATCCATTAATGCCCTGGCTTCATTCGGTCACTCTGTGTGTGTGCTTTTACCACCGC	4242
3661	Db	ATGATCCATTAATGCCCTGGCTTCATTCGGTCACTCTGTGTGTGTGCTTTTACCACCGC	3720
4243	Qy	GTCATCTCTGAGGAAGTCACTTTCCACCTCTACTGATCCCAAGTGAAGTCTCTCGATCCG	4302
3721	Db	GTCATCTCTGAGGAAGTCACTTTCCACCTCTACTGATCCCAAGTGAAGTCTCTCGATCCG	3780
4303	Qy	AAGGAACTGAGCTCTGCTATCGAAGCCCTGGAGAACCCAGCTGTTCTCGGAGTTCTAC	4362
3781	Db	AAGGAACTGAGCTCTGCTATCGAAGCCCTGGAGAACCCAGCTGTTCTCGGAGTTCTAC	3840
4363	Qy	GTTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGATGAGACTCTG	4422
3841	Db	GTTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGATGAGACTCTG	3900
4423	Qy	GTGTGGGAGCCCTTGGTGAACACAGAGATCTCATGCTGCAACTCTCTGATCCCTCCA	4482
3901	Db	GTGTGGGAGCCCTTGGTGAACACAGAGATCTCATGCTGCAACTCTCTGATCCCTCCA	3960
4483	Qy	GCCCGCATGCCGTACCTTCACTCTGGATGCCCGCAGTGTCTGCACTTTGTGGACCA	4542
3961	Db	GCCCGCATGCCGTACCTTCACTCTGGATGCCCGCAGTGTCTGCACTTTGTGGACCA	4020
4543	Qy	TATCGAGACAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTCTTGGACAAACTGCAT	4602
4021	Db	TATCGAGACAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTCTTGGACAAACTGCAT	4080

QY 301 GTCTTGCAGGCCCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTAT 360
 Db 331 GTCTTGCAGGCCCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTTACCACTCCCTCTAT 390
 QY 361 CGGCGTCTATCTGTAGGTGCCCCGGAATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
 Db 391 CGGCGTCTATCTGTAGGTGCCCCGGAATTTATAAACTGGGTTCCGAATGCTGAATAAGA 450
 QY 421 GACGGTAAAGCAAGGCAAGGACAGCACTGTCTCTCCCTGCTGTATACCTCACCAC 480
 Db 451 GACGGTAAAGCAAGGCAAGGACAGCACTGTCTCTCCCTGCTGTATACCTCACCAC 510
 QY 481 CTGGGAACATCCCCAGACACCTCTTAATCCGGGACAGAGATGGCTGGCGAGCCTGG 540
 Db 511 CTGGGAACATCCCCAGACACCTCTTAATCCGGGACAGAGATGGCTGGCGAGCCTGG 570
 QY 541 GGCCTGCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 600
 Db 571 GGCCTGCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 630
 QY 601 CTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTGGGAGAGACACCCGCTCAGCCAGAG 660
 Db 631 CTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTGGGAGAGACACCCGCTCAGCCAGAG 690
 QY 661 AAGACGAGTGGCATGGAGGTGGCTCGTACCTGGTGGCTCAGTATGGGAGAGCGGGCC 720
 Db 691 AAGACGAGTGGCATGGAGGTGGCTCGTACCTGGTGGCTCAGTATGGGAGAGCGGGCC 750
 QY 721 TGGGACCTTAGCCCTCATACCTGGGAGAGATGGGCTGAGTCACTGTGGCCCAAGCC 780
 Db 751 TGGGACCTTAGCCCTCATACCTGGGAGAGATGGGCTGAGTCACTGTGGCCCAAGCC 810
 QY 781 CAGGAAGGGGAGGCACTCTCCCTCATCTCCCTACAGCCCAAGTGAACCCACCTGGGG 840
 Db 811 CAGGAAGGGGAGGCACTCTCCCTCATCTCCCTACAGCCCAAGTGAACCCACCTGGGG 870
 QY 841 TCTCCAGGCAACCCACCTCCACCGCAGTGTCTAAATGCCCTGGATCCATGAATGGCGGG 900
 Db 871 TCTCCAGGCAACCCACCTCCACCGCAGTGTCTAAATGCCCTGGATCCATGAATGGCGGG 930
 QY 901 GGGTGACCCAGGCTCAGAGAGAGGGTTTGAACAGCTGCTGACACATCTGAGAGC 960
 Db 931 GGGTGACCCAGGCTCAGAGAGAGGGTTTGAACAGCTGCTGACACATCTGAGAGC 990
 QY 961 CGCTGGAGAGAAATCTCTGGCTCACTCTCTACCAAGCTTTCCAGCTCCCGACCACT 1020
 Db 991 CGCTGGAGAGAAATCTCTGGCTCACTCTCTACCAAGCTTTCCAGCTCCCGACCACT 1050
 QY 1021 GAGTCTCCAAAGCAGGAGTCAACCAACGCCCCACATCCACAGCAGTGTGGGAGCTGG 1080
 Db 1051 GAGTCTCCAAAGCAGGAGTCAACCAACGCCCCACATCCACAGCAGTGTGGGAGCTGG 1110
 QY 1081 GGATCCCCACCTCAGCCAGCTAGCAACCCAGAGAGAGGCTCCTGGGACCCCAATGG 1140
 Db 1111 GGATCCCCACCTCAGCCAGCTAGCAACCCAGAGAGAGGCTCCTGGGACCCCAATGG 1170
 QY 1141 CCTCTGGATGAACGTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAA 1200
 Db 1171 CCTCTGGATGAACGTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAA 1230
 QY 1201 TCAGAGAAAGGAGGCCCCCATGGGAGCGGTGGTAGGAACGCCCCACAGGGCGACACC 1260
 Db 1231 TCAGAGAAAGGAGGCCCCCATGGGAGCGGTGGTAGGAACGCCCCACAGGGCGACACC 1290
 QY 1261 AGCCTACGCCCCACACCAACCCATCGGAGCTTCTGTGAGAGAGAGCCTCTGTCCACA 1320
 Db 1291 AGCCTACGCCCCACACCAACCCATCGGAGCCTTCTGTGAGAGAGAGCCTCTGTCCACA 1350
 QY 1321 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTCACAGAGTGTCTTCTACAAAGA 1380
 Db 1351 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTCACAGAGTGTCTTCTACAAAGA 1410

QY 1381 CCTACCCCAAGCAAGATCCCTCTGTCAAGAGAACTGGCTGTATATGTGAGAG 1440
 Db 1411 TCTCACCCCAAGCAAGATCCCTCTGTCAAGAGAACTGGCTGTATATGTGAGAG 1470
 QY 1441 AATCGAGGACATTTAATTTGAGATCAGAGACTTATTTGGCCAGGCTCGATACCAAGAA 1500
 Db 1471 AATCGAGGACATTTAATTTGAGATCAGAGACTTATTTGGCCAGGCTCGATACCAAGAA 1530
 QY 1501 CTTCCATAGTCATCTCAGGGGCTCTGGAATTTGGAAAGTCAACACTGGCCAGGAG 1560
 Db 1531 CTTCCATAGTCATCTCAGGGGCTCTGGAATTTGGAAAGTCAACACTGGCCAGGAG 1590
 QY 1561 GTGAAGAAAGCTGGGGAGAGGCGAGCTGTATGGGACCGCTTCAGAGATGTCTTAC 1620
 Db 1591 GTGAAGAAAGCTGGGGAGAGGCGAGCTGTATGGGACCGCTTCAGAGATGTCTTAC 1650
 QY 1621 TTCAGCTCAGAGAGCTGGCCAGTCCAAGTGGTGTAGTCTCGCTGAGCTCATCGGAAA 1680
 Db 1651 TTCAGCTCAGAGAGCTGGCCAGTCCAAGTGGTGTAGTCTCGCTGAGCTCATCGGAAA 1710
 QY 1681 GATGGGACAGCCACTCCGGCTCCCATTAGACAGATCTCTGTAGGCCAGAGCGGTCTC 1740
 Db 1711 GATGGGACAGCCACTCCGGCTCCCATTAGACAGATCTCTGTAGGCCAGAGCGGTCTC 1770
 QY 1741 TTCATCTCGATGGTGTAGTAGGACAGAGTGGTCTTTCAGGAGCGGAGTCTGAGCTC 1800
 Db 1771 TTCATCTCGATGGTGTAGTAGGACAGAGTGGTCTTTCAGGAGCGGAGTCTGAGCTC 1830
 QY 1801 TGTCTGCACTGGAGGACGCCACAGCGGGGATGCACTGTGGGAGTTCGTGGGAAA 1860
 Db 1831 TGTCTGCACTGGAGGACGCCACAGCGGGGATGCACTGTGGGAGTTCGTGGGAAA 1890
 QY 1861 ACTATACTTCCGAGGCACTCTTCTGTATCAGGCTCGGACCAAGCTCTGAGAACCTC 1920
 Db 1891 ACTATACTTCCGAGGCACTCTTCTGTATCAGGCTCGGACCAAGCTCTGAGAACCTC 1950
 QY 1921 ATTCTCTTTGGAGCAGGACGTTGGTGTAGGTCCTGGGTTCTGTAGTTCAGCAGG 1980
 Db 1951 ATTCTCTTTGGAGCAGGACGTTGGTGTAGGTCCTGGGTTCTGTAGTTCAGCAGG 2010
 QY 1981 AAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGACCTTTAGTTG 2040
 Db 2011 AAGGAATATTTCTACAGATATTTCTACAGATGAAGGCAAGCAATTAGACCTTTAGTTG 2070
 QY 2041 GTCAATCAAAAGAGCTCTGGCCCTGTCTGTGTGGCTGTGGTGTCTGTGGCTGGCC 2100
 Db 2071 GTCAATCAAAAGAGCTCTGGCCCTGTCTGTGTGGCTGTGGTGTCTGTGGCTGGCC 2130
 QY 2101 TGCACTTGCCTGATCAGCAGATGAAGCGGAAAGGAAAACTCACACTGACTTCCAAGACC 2160
 Db 2131 TGCACTTGCCTGATCAGCAGATGAAGCGGAAAGGAAAACTCACACTGACTTCCAAGACC 2190
 QY 2161 ACCAAGCCCTGTCTACATTTACCTTGGCCAGGCTCTCCAAGCTCAGCCATTTGGACCC 2220
 Db 2191 ACCAAGCCCTGTCTGTCTACATTTACCTTGGCCAGGCTCTCCAAGCTCAGCCATTTGGACCC 2250
 QY 2221 CAGCTCAGAGACTCTGTCTCTGTCTGTAGGCGCATCTGGCAAAAAGAACCTTTTC 2280
 Db 2251 CAGCTCAGAGACTCTGTCTCTGTCTGTAGGCGCATCTGGCAAAAAGAACCTTTTC 2310
 QY 2281 AGTCCAGATGACCTCAGGAAGCATGGTGTAGATGGGCGCATCATCTCCAACCTTTGAAG 2340
 Db 2311 AGTCCAGATGACCTCAGGAAGCATGGTGTAGATGGGCGCATCATCTCCAACCTTTGAAG 2370
 QY 2341 ATGGGTATTTTCAAGAGCAACCCATCTCCTGTAGCTACAGCTTCACTTCCCTCTGTTC 2400
 Db 2371 ATGGGTATTTTCAAGAGCAACCCATCTCCTGTAGCTACAGCTTCACTTCCCTCTGTTC 2430
 QY 2401 CAAGAGTCTTTTGCAGCAATGTCTTATGTCTTGGAGGATGAGAAGGGGAGAGTAACAT 2460
 Db 2431 CAAGAGTCTTTTGCAGCAATGTCTTATGTCTTGGAGGATGAGAAGGGGAGAGTAACAT 2490
 QY 2461 TCTAATTTGCATCATAGATTTTGGAAAAAGCGCTAGAGCATATATGAATATATGCTGTTT 2520

[illegible]

LOCUS AX089761 4422 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 1 from Patent WO0116170.
ACCESSION AX089761
VERSION AX089761.1 GI:13443933
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Reed, J.C.
TITLE Card proteins involved in cell death regulation
JOURNAL Patent: WO 0116170-A 1 08-MAR-2001;
The Burnham Institute (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS
1..4422
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34899.1"
/db_xref="GI:13443934"
/translation="MAGGAWGLACYLEFKKELKBFOLLILANKAHSRSSGETPAQ
PEKTSMEVASYLVIAQYGRQAWDLALHTWEQWGLRSLCAQAGAGHSPSPFSPSE
PHLGSPQSTAVLMPWTHLPAGCTQGSERRVLRPLDPTSGRWREISASLLYQAL
RSSPHESPSQSPNAPTAVLGSWGPSPOPLAPREOEAPGTOWPLDETSGIYVTE
IREREREKSEKRPAAVVGTPPQNAHTSLQPHHPWEPVSRESLCSWPKNEPNO
KFTQLLLQRPSPQDPYVEENRGLHIEIRDLFGPLDQTPRIVILQG
APRIQLSPERLLFLDGDVDPGWLVFPSESLCHWSQPOPADALLGSLGKTLTP
AAGIGKSTLRAVKEAGRGQYGRFQHFVYFSCRELAQSKVSLAEILIGKDTATP
EASFLLTATLWNLIPLEQARWVEVLGFSESRKEVYFYFDDEROAIRAFRLVK
SNKELWALCLVPMWNLATCLMOWMRKEKLTSTKTTTLLHVLQAALAOALPLP
QURLDLSLAEGIWQKTLFSPDDLKRGIDGAIISTFLMGILQHPPLPSYSIHL
CFQEPFAASYVLEDEKRGKSNCLIDLEKTEAVGIHGLFGASTFELLGLLSDG
EREMENIFRCLSQNLMQWVPSLQILLPHSLHSLCHLYETRNKTFILTMVAHFEE
MGCVETDMELLVCTFCIKFSRHKVKKLQIEGRQHRSTWSPVTLFRWVPVTDAYWQ
ILFVLVTRNLIKELDLSGNSLSHSAVSLCKLTRRCLLETLLRAGGLTAEDCKD
LAFGLRANOTLFELDSLNVLDAGAKHLQBLRPSCKLQBLQVSCGLTSDCCODL
ASVLSAPSLKELDLQNNLDVGVRLLCGLRHHPACKLIRGLDOITLSDMRQLR
ALEQKPOLLIISRRKPSVMTPTGELTGEMNSLTKRQRLGSERAASHVAQANLK
LIDVSKIFPIAIAEESSEVVPVELLCPSPASQDHLTKPLGDDDDFWGTPGVAT
EVDVKEKNLYRVHFPVAGSVRWPNTPGLCFVMEAVTEVEFCVMDQFLGEINPQHSWM
VAGPLIDIKAEPAVGAHVLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARVSLHH
IVLENPSPSLPVLKWIHNLARFIPVTSVLLYHRVHPVEVTHLYLIPSDCSIRKA
IDLEMKFQFVRHKKPPPLTPYMGCRYTVSGSGSNLSEILPKLELCLVRSRGEQLF
SBFYVGLSGIRLQVKDKDETLVWEALVPGDLMPATLILPPARIAVPSPLDAPQL
LHFVDYRQLIARVTSVSVLKLHGQVLSQEQYERVLAEINTRPSQMRKLFSLSQSW
DRKCKDGLYQALKETHPLHIMELWEKSKKGLLPLSS"
BASE COUNT 1043 a 1224 c 1240 g 915 t
ORIGIN
Query Match 76.2%; Score 4148; DB 6; Length 4422;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;
QY 523 ATGGCTGGCGGAGCCTGGGCGCCCTGGCCCTGTTACTTGGAGTTCCTGAAGAGGAGGAG 582
DB 1 ATGGCTGGCGGAGCCTGGGCGCCCTGGCCCTGTTACTTGGAGTTCCTGAAGAGGAGGAG 60
QY 583 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACCTCCAGGAGCTCTTCGGGTGAG 642
DB 61 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACCTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCAGAGAACGAGTGGCATGGAGTGGCCCTCGTACCTGGTGGCTCAG 702
DB 121 ACACCCGCTCAGCAGAGAACGAGTGGCATGGAGTGGCCCTCGTACCTGGTGGCTCAG 180
QY 703 TATGGGAGGAGCGGCGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCGTGGG 762
DB 181 TATGGGAGGAGCGGCGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCGTGGG 240

QY 763 TCACCTGTGGCGCCCAAGCCAGAGAGGGGAGGCCACCTCTCCCTCATTTCCCTTACAGCCCA 822
DB 241 TCACCTGTGGCGCCCAAGCCAGAGAGGGGAGGCCACCTCTCCCTCATTTCCCTTACAGCCCA 300
QY 823 AGTGAACCCCACTCTGGGTCTCCAGCAACCCACCTCCACCGCAGTGTCTAATGACCTGG 882
DB 301 AGTGAACCCCACTCTGGGTCTCCAGCAACCCACCTCCACCGCAGTGTCTAATGACCTGG 360
QY 883 ATCCATGAATTTGCCCGCGGGGTGACCCAGGSGCTCAGAGAGAGGGTTTGTGACAGCTG 942
DB 361 ATCCATGAATTTGCCCGCGGGGTGACCCAGGSGCTCAGAGAGAGGGTTTGTGACAGCTG 420
QY 943 CTTGACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTCTT 1002
DB 421 CTTGACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTCTT 480
QY 1003 CCAAGCTCCAGAGCAATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCATCCACA 1062
DB 481 CCAAGCTCCAGAGCAATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCATCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGCAGGAG 1122
DB 541 GCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGCAGGAG 600
QY 1123 GCTCTGGGACCCAATGGCTCTGGATGAACCTGAGAAATTTATACACAGAAATCAGA 1182
DB 601 GCTCTGGGACCCAATGGCTCTGGATGAACCTGAGAAATTTATACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCGCCCATGGGACGGTGTGAGAAAG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCGCCCATGGGACGGTGTGAGAAAG 720
QY 1243 CCCCCACAGGCGCACACAGCGCTACAGGCCACACACACCTAGAGAGGCTTCTGTGAGA 1302
DB 721 CCCCCACAGGCGCACACAGCGCTACAGGCCACACACACCTAGAGAGGCTTCTGTGAGA 780
QY 1303 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATAGGATTTTAAACAAAAATTCACACAG 1362
DB 781 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATAGGATTTTAAACAAAAATTCACACAG 840
QY 1363 CTGCTACTTCTACAAAGACCTACCCAGAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG 1422
DB 841 CTGCTACTTCTACAAAGACCTACCCAGAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG 900
QY 1423 CTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 1482
DB 901 CTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACCTCGCATAGTTCATCTGAGGGGGCTGCTGGAATTTGGGAAG 1542
DB 961 GGCCTGGATACCAAGAACCTCGCATAGTTCATCTGAGGGGGCTGCTGGAATTTGGGAAG 1020
QY 1543 TCACACCTGGCCAGGAGGTGAAGAGCCTGGGGAGAGGGCAGCTGTATGGGAGCCGC 1602
DB 1021 TCAACACCTGGCCAGGAGGTGAAGAGCCTGGGGAGAGGGCAGCTGTATGGGAGCCGC 1080
QY 1603 TTTCAGACATGCTCTTACTTTCAGCTGCAGAGAGCTGGCCCAAGTCCAAAGTGTGAGTCTC 1662
DB 1081 TTTCAGACATGCTCTTACTTTCAGCTGCAGAGAGCTGGCCCAAGTCCAAAGTGTGAGTCTC 1140
QY 1663 GGTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGGCTCCCATTTAGACAGATCCTGTCT 1722
DB 1141 GGTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGGCTCCCATTTAGACAGATCCTGTCT 1200
QY 1723 AGGCCAGAGCGGCTGCTTTCATCTCGATGTGTAGATGAGCCAGGATGGGTCTTGGAG 1782
DB 1201 AGGCCAGAGCGGCTGCTTTCATCTCGATGTGTAGATGAGCCAGGATGGGTCTTGGAG 1260
QY 1783 GAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCACAGCCGCGGATGCACCTGCTG 1842
DB 1261 GAGCCGAGTTCTGAGCTGTCTGCACTGGAGCCACAGCCGCGGATGCACCTGCTG 1320
QY 1843 GGCAGTTTCTGGGGAAAAACTATATCTTCCGAGGCACTCTTCTGTATCAGCGCTCGGACC 1902

Db		1321	GGCAGTTTGTGGGAAAACTA	
Qy		1903	ACAGCTCTGGAGAACTCATT	
Db		1381	ACAGCTCTGCAGAACCTCAT	
Qy		1963	TTCTCTGAGTCCAGCAGGA	
Db		1441	TTCTCTGAGTCCAGCAGGA	
Qy		2023	ATTAGAGCCCTTTAGGTGG	
Db		1501	ATTAGAGCCCTTTAGGTGG	
Qy		2083	TGGGTGTCTGGCTGGCTGC	
Db		1561	TGGGTGTCTGGCTGGCTGC	
Qy		2143	ACACTGACTTCCAAAGACC	
Db		1621	ACACTGACTTCCAAAGACC	
Qy		2203	GCTCAGCCATTGGGACCC	
Db		1681	GCTCAGCCATTGGGACCC	
Qy		2263	CAAAAAAGACCCCTTTCA	
Db		1741	CAAAAAAGACCCCTTTCA	
Qy		2323	ATCTCAGCCCTTTCAAGAT	
Db		1801	ATCTCAGCCCTTTCAAGAT	
Qy		2383	TTCAATTCACCTCTGTTT	
Db		1861	TTCAATTCACCTCTGTTT	
Qy		2443	AAGGGAGAGGTAACATTC	
Db		1921	AAGGGAGAGGTAACATTC	
Qy		2503	GGAAATACATGGCCCTGT	
Db		1981	GGAAATACATGGCCCTGT	
Qy		2563	GAGGGGAGAGAGAGATGG	
Db		2041	GAGGGGAGAGAGAGATGG	
Qy		2623	ATGCAGTGGGTCCCGTCC	
Db		2101	ATGCAGTGGGTCCCGTCC	
Qy		2683	TGCTTGTACGAGCTCGGA	
Db		2161	TGCTTGTACGAGCTCGGA	
Qy		2743	ATGGGCATGTGTAGAAAC	
Db		2221	ATGGGCATGTGTAGAAAC	
Qy		2803	AGCCGCCAGTGAAGCTTC	
Db		2281	AGCCGCCAGTGAAGCTTC	
Qy		2863	CCACCATGGTAGTCTGTT	
Db		2341	CCACCATGGTAGTCTGTT	
Qy		2923	TTCTCCGTCTCAAGGTTC	

Db		2401	TTCTCCGTCTCAAGGTTC	
Qy		2983	AGCCACTCTGCAGTGAAG	
Db		2461	AGCCACTCTGCAGTGAAG	
Qy		3043	ACCTCTGCGGTTGGCTGT	
Db		2521	ACCTCTGCGGTTGGCTGT	
Qy		3103	CTGAGAGCAACACGACCT	
Db		2581	CTGAGAGCAACACGACCT	
Qy		3163	GGAGCCAAACACCTTTG	
Db		2641	GGAGCCAAACACCTTTG	
Qy		3223	CTGGTCAGCTGTGGCTCA	
Db		2701	CTGGTCAGCTGTGGCTCA	
Qy		3283	AGCCCCAGCCTGAAGGAG	
Db		2761	AGCCCCAGCCTGAAGGAG	
Qy		3343	CTGCTCTGTGAGGGGCT	
Db		2821	CTGCTCTGTGAGGGGCT	
Qy		3403	ACAACTCTGAGTGATGAG	
Db		2881	ACAACTCTGAGTGATGAG	
Qy		3463	CTGCTCATCTTCAGCAGA	
Db		2941	CTGCTCATCTTCAGCAGA	
Qy		3523	GGAGAGATGAGTAAATAG	
Db		3001	GGAGAGATGAGTAAATAG	
Qy		3583	GCTTCCCATGTTGCTCAG	
Db		3061	GCTTCCCATGTTGCTCAG	
Qy		3643	GCTGAGATGAGAGGAAAG	
Db		3121	GCTGAGATGAGAGGAAAG	
Qy		3703	TCTCTGCTCTCAAGGGGA	
Db		3181	TCTCTGCTCTCAAGGGGA	
Qy		3763	GGCCCAACGGGCGCTGT	
Db		3241	GGCCCAACGGGCGCTGT	
Qy		3823	CATTTCCCTGTAGCTGG	
Db		3301	CATTTCCCTGTAGCTGG	
Qy		3883	GNAGCGGTGACCGTTGAG	
Db		3361	GNAGCGGTGACCGTTGAG	
Qy		3943	CCACAGCAGCTGGATGGT	
Db		3421	CCACAGCAGCTGGATGGT	
Qy		4003	GTGGAAGCTGTGCACCT	
Db		3481	GTGGAAGCTGTGCACCT	

Db 121 ACACCGCTCAGCAGAGAAAGCAGTGGCATGGAGTGGCTCTGTAACCTGTGGCTCAG 180
QY 703 TATGGGAGCAGCGGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTTGAGG 762
Db 181 TATGGGAGCAGCGGCTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTTGAGG 240
QY 763 TCACTGTGGCCCAAGCCCAAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCA 822
Db 241 TCACTGTGGCCCAAGCCCAAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCACAACCACTCCACCGCAGTGTCTAATGCCCTGG 882
Db 301 AGTGAACCCCACTGGGGTCTCCAGCACAACCACTCCACCGCAGTGTCTAATGCCCTGG 360
QY 883 ATCCATGAATTGCGCGGGGTGSCACCCAGGGCTCAGAGAGAGGGTGTGAGACAGCTG 942
Db 361 ATCCATGAATTGCGCGGGGTGSCACCCAGGGCTCAGAGAGAGGGTGTGAGACAGCTG 420
QY 943 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCCTCTACCAAGCTCTT 1002
Db 421 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAAGCACAATGATCTCAAGCAGAGTCAACCAAGCGCCCACTCCACA 1062
Db 481 CCAAGCTCCCAAGCACAATGATCTCAAGCAGAGTCAACCAAGCGCCCACTCCACA 540
QY 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGCCAGAGCAGGAG 1122
Db 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGCCAGAGCAGGAG 600
QY 1123 GCTCTGGGACCAATGGGCTCTGGATGAAACGTCAAGAAATTTACTACACAGAAATCAGA 1182
Db 601 GCTCTGGGACCAATGGGCTCTGGATGAAACGTCAAGAAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGCCCAATGGGCGCGGTGTAGGAACG 1242
Db 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGCCCAATGGGCGCGGTGTAGGAACG 720
QY 1243 CCCCCAGCGCACACAGCTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA 1302
Db 721 CCCCCAGCGCACACAGCTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAAATTCACACAG 1362
Db 781 GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAAATTCACACAG 840
QY 1363 CTGCTACTTCTACAAAGACTCAACCCAGAGCAGCAAGATCCCTGTGTAAGAGAGCTGG 1422
Db 841 CTGCTACTTCTACAAAGACTCAACCCAGAGCAGCAAGATCCCTGTGTAAGAGAGCTGG 900
QY 1423 CTTGATTATGTGGAGAGATCGAGACATTTAATTGAGATCAGACATTTTGGCCCA 1482
Db 901 CTTGATTATGTGGAGAGATCGAGACATTTAATTGAGATCAGACATTTTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACTCGCATAGTCATCTCAGGGGCTCTCGAAATTTGGGAAG 1542
Db 961 GGCCTGGATACCAAGAACTCGCATAGTCATCTCAGGGGCTCTCGAAATTTGGGAAG 1020
QY 1543 TCAACACTGGCCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATGGGACCGC 1602
Db 1021 TCAACACTGGCCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATGGGACCGC 1080
QY 1603 TTCCAGCATGCTTCTACTTCAGTCCAGAGCTGGCCAGTCCAAAGTGGTGTGCTC 1662
Db 1081 TTCCAGCATGCTTCTACTTCAGTCCAGAGCTGGCCAGTCCAAAGTGGTGTGCTC 1140
QY 1663 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCTGTCT 1722
Db 1141 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCTGTCT 1200
QY 1723 AGGCCAGAGCGGTCTCTTCACTCCGATGCTGTAGTAGAGCCAGATGGGCTTTGACG 1782
Db 1201 AGGCCAGAGCGGTCTCTTCACTCCGATGCTGTAGTAGAGCCAGATGGGCTTTGACG 1260

QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG 1842
Db 1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAAATATATCTCCGAGGATCTCTCTGATCAGCGCTGGACC 1902
Db 1321 GGCAGTTTGTGGGAAAAATATATCTCCGAGGATCTCTCTGATCAGCGCTGGACC 1380
QY 1903 ACAGCTCTGCAGAACCTCATTTCTTTGGAGAGGACAGCTGGGTAGAGTCTGGGG 1962
Db 1381 ACAGCTCTGCAGAACCTCATTTCTTTGGAGAGGACAGCTGGGTAGAGTCTGGGG 1440
QY 1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 1500
QY 2023 ATTAGAGCTTTAGGTTGGTCAAATCAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
Db 1501 ATTAGAGCTTTAGGTTGGTCAAATCAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560
QY 2083 TGGGTCTCTGGCTGGCTGCACCTTGCCTGATGACAGAGATGAAGCGAAGAAAACTC 2142
Db 1561 TGGGTCTCTGGCTGGCTGCACCTTGCCTGATGACAGAGATGAAGCGAAGAAAACTC 1620
QY 2143 ACAGTGAATTTCAAAGACCAACCAACCACTCTGTCTACATTACCTTGGCCAGGCTCTCAA 2202
Db 1621 ACAGTGAATTTCAAAGACCAACCAACCACTCTGTCTACATTACCTTGGCCAGGCTCTCAA 1680
QY 2203 GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGTCTGAGGCACTCGG 2262
Db 1681 GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGTCTGAGGCACTCGG 1740
QY 2263 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGTTAGATGGGCGCATC 2322
Db 1741 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGTTAGATGGGCGCATC 1800
QY 2323 ATCTCCACCTTCTGAAGATGGGTATTTCTCAAGAGCACCCCATCTCTGAGCTACAGC 2382
Db 1801 ATCTCCACCTTCTGAAGATGGGTATTTCTCAAGAGCACCCCATCTCTGAGCTACAGC 1860
QY 2383 TTCAATCACTCTGTCTTCAAAGAGTTCTTTCAGCAATGTCTATGTCTTGGAGATGAG 2442
Db 1861 TTCAATCACTCTGTCTTCAAAGAGTTCTTTCAGCAATGTCTATGTCTTGGAGATGAG 1920
QY 2443 AAGGGAGAGGTAAACATTTCTAATTTGCAATCATAGATTTGGAAGACCCCTAGAACATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTCTAATTTGCAATCATAGATTTGGAAGACCCCTAGAACATAT 1980
QY 2503 GGAATACATGGCTCTTTGGGCGCATCAACCAACCGTTTCTTATTTGGGCTGTAAAGTAT 2562
Db 1981 GGAATACATGGCTCTTTGGGCGCATCAACCAACCGTTTCTTATTTGGGCTGTAAAGTAT 2040
QY 2563 GAGGGGAGAGAGATGGAGAACTTTTCACTCCGGCTGTCTCAGGGGAGAACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAACTTTTCACTCCGGCTGTCTCAGGGGAGAACCTG 2100
QY 2623 ATGAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTCCCTCCAC 2682
Db 2101 ATGAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTCCCTCCAC 2160
QY 2683 TGCTTGTACGAGACTCGGAACAAACCGTTCTTGACACAGTGTAGTGGCCATTTCAAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAACAAACCGTTCTTGACACAGTGTAGTGGCCATTTCAAGAA 2220
QY 2743 ATGGCATGTGTGTAGAAACAGACATGGAGCTTTTAGTGTGCACTTTCTGCAATTAATTC 2802
Db 2221 ATGGCATGTGTGTAGAAACAGACATGGAGCTTTTAGTGTGCACTTTCTGCAATTAATTC 2280
QY 2803 AGCCGCCAGGTGAAGAACTTTCAGCTGATTTAGGGCAGGACAGACATCAACATGGAGC 2862
Db 2281 AGCCGCCAGGTGAAGAACTTTCAGCTGATTTAGGGCAGGACAGACATCAACATGGAGC 2340

Qy	2863	CCACACATGTGATGTCCTGTTTCAGGTGGGTCCAGATCTACAGATGCCCTATTTGGCAGATTCTC	2922
Db	2341	CCCAACCATGTGTAGTCTCTGTTTCAGGTGGGTCCCACTCTACAGATGCCCTATTTGGCAGATTCTC	2400
Qy	2923	TTCTCCGTCCTCAAGGTTCACAGAAACCTGAAGGAGCTGGACCTAAGTGTGAACCTCGCTG	2982
Db	2401	TTCTCCGTCCTCAAGGTTCACAGAAACCTGAAGGAGCTGGACCTAAGTGTGAACCTCGCTG	2460
Qy	2983	AGGCACCTCTGCAAGTGAAGAGTCTTTGTTAAGACCCCTGAGAGCCCTCGCTGCCTCTCTGGAG	3042
Db	2461	AGGCACCTCTGCAAGTGAAGAGTCTTTGTTAAGACCCCTGAGAGCCCTCGCTGCCTCTCTGGAG	2520
Qy	3043	ACCTCGGGTTGGCTGTGGCTCTACAGCTGAGGACTGCAAGGACCTTGCTTGGTGGG	3102
Db	2521	ACCTCGGGTTGGCTGTGGCTCTACAGCTGAGGACTGCAAGGACCTTGCTTGGTGGG	2580
Qy	3103	CTGAGAGCCAAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT	3162
Db	2581	CTGAGAGCCAAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT	2640
Qy	3163	GGAGCCAAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCAG	3222
Db	2641	GGAGCCAAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCAG	2700
Qy	3223	CTGCTGAGCTGTGGCCTCAAGTCTGAGTGTGCCAGGACCTGGCCTCTGTGTAGTGGC	3282
Db	2701	CTGCTGAGCTGTGGCCTCAAGTCTGAGTGTGCCAGGACCTGGCCTCTGTGTAGTGGC	2760
Qy	3283	AGCCCCAGCCTGAAGGAGCTAGACCTCCAGCAGAAACCTGGATGACGTTGGCGTGCGA	3342
Db	2761	AGCCCCAGCCTGAAGGAGCTAGACCTCCAGCAGAAACCTGGATGACGTTGGCGTGCGA	2820
Qy	3343	CTGCTCTGTGAGGGGCTCAGGCATCTCCCTCTGCAAACTCATACGCCCTGGGGCTGGACCAG	3402
Db	2821	CTGCTCTGTGAGGGGCTCAGGCATCTCCCTCTGCAAACTCATACGCCCTGGGGCTGGACCAG	2880
Qy	3403	ACAACTCTGAGTGTATGATGAGCGCAGAACTGAGGGCCCTGGAGCAGGAGAAACCTCTCAG	3462
Db	2881	ACAACTCTGAGTGTATGATGAGCGCAGAACTGAGGGCCCTGGAGCAGGAGAAACCTCTCAG	2940
Qy	3463	CTGCTCATCTTCACGACGCGAAACCAAGTGTGATGATCCCTACTGAGGGCCCTGGATACG	3522
Db	2941	CTGCTCATCTTCACGACGCGAAACCAAGTGTGATGATCCCTACTGAGGGCCCTGGATACG	3000
Qy	3523	GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
Db	3001	GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3060
Qy	3583	GCTTCCCATGTTGCTCAGGCTAAATCTCAAACTCTCGACGTGAGCAAGATCTTCCCAATT	3642
Db	3061	GCTTCCCATGTTGCTCAGGCTAAATCTCAAACTCTCGACGTGAGCAAGATCTTCCCAATT	3120
Qy	3643	GCTGAGATTGACAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTTGCGGTGCCT	3702
Db	3121	GCTGAGATTGACAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTTGCGGTGCCT	3180
Qy	3703	TCTCTGTGCTCTCAAGGGGACCTGCATACGAAGCCTTTTGGGGACTGACGATGACTTCTGG	3762
Db	3181	TCTCTGTGCTCTCAAGGGGACCTGCATACGAAGCCTTTTGGGGACTGACGATGACTTCTGG	3240
Qy	3763	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGCAAAAGAAAGAACTTGTACCGAGTT	3822
Db	3241	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGCAAAAGAAAGAACTTGTACCGAGTT	3300
Qy	3823	CACHTCCCTGTAGCTGGCTCCTACGGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
Db	3301	CACHTCCCTGTAGCTGGCTCCTACGGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGA	3360
Qy	3883	GAAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3942
Db	3361	GAAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3420
Qy	3943	CCACAGCACAGCTGGATGGTGGCAGGGCCCTCTGCTGGAACATCAAGGCTGAGCCTGGAGCT	4002

RESULT 10
AF229061
LOCUS
DEFINITION

[illegible]

Db 1141 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1200
Qy 1723 AGCCAGAGCGCTGTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGCTTTGCAG 1782
Db 1201 AGGCCAGAGCGCTGTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGCTTTGCAG 1260
Qy 1783 GAGCCGAGTTCTGAGCTGTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGTCTG 1842
Db 1261 GAGCCGAGTTCTGAGCTGTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGTCTG 1320
Qy 1843 GGCAGTTTGTGGGAAATATATCTTCCGAGGATCTTCTCTGATCAGCGCTCGGACC 1902
Db 1321 GGCAGTTTGTGGGAAATATATCTTCCGAGGATCTTCTCTGATCAGCGCTCGGACC 1380
Qy 1903 ACAGCTCTGCAGAACTCATCTCTTCTTGGAGCAGGACAGTGTGGGTAGAGCTCTGGGG 1962
Db 1381 ACAGCTCTGCAGAACTCATCTTCTTCTTGGAGCAGGACAGTGTGGGTAGAGCTCTGGGG 1440
Qy 1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCA 1500
Qy 2023 ATTAGAGCTTTAGTGTGTGTCAAATCAACAAGAGCTCTGGGCCCTGTGTCTGTGCC 2082
Db 1501 ATTAGAGCTTTAGTGTGTGTCAAATCAACAAGAGCTCTGGGCCCTGTGTCTGTGCC 1560
Qy 2083 TGGGTGTCTGGCTGGCTGCATCTTGCCTGTGATGACAGATGAAGCGGAAGAAAATCT 2142
Db 1561 TGGGTGTCTGGCTGGCTGCATCTTGCCTGTGATGACAGATGAAGCGGAAGAAAATCT 1620
Qy 2143 ACACAGCTTTCAAGACCAACCAACCCCTCTGTCTACATTTACCTTGCOCAGGCTCTCCAA 2202
Db 1621 ACACAGCTTTCAAGACCAACCAACCCCTCTGTCTACATTTACCTTGCOCAGGCTCTCCAA 1680
Qy 2203 GCTCAGCAATGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGTGGGCACTCTGG 2262
Db 1681 GCTCAGCAATGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGTGGGCACTCTGG 1740
Qy 2263 CAAAAAGACCCCTTTTCACTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC 2322
Db 1741 CAAAAAGACCCCTTTTCACTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC 1800
Qy 2323 ATCTCCACCTTTTGAAGATGGGTATTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
Db 1801 ATCTCCACCTTTTGAAGATGGGTATTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860
Qy 2383 TTCAATCACTCTGTTTCAAGAGTCTTTGAGCAATGTCTATGTCTTGAGAGATGAG 2442
Db 1861 TTCAATCACTCTGTTTCAAGAGTCTTTGAGCAATGTCTATGTCTTGAGAGATGAG 1920
Qy 2443 AAGGGGAGAGTAAACATTTCTAATTCATCATAGATTTTGGAAAGACGCTAGAAGCATAT 2502
Db 1921 AAGGGGAGAGTAAACATTTCTAATTCATCATAGATTTTGGAAAGACGCTAGAAGCATAT 1980
Qy 2503 GGAATACATGGCTGTGTGGGCGATCAACACACAGTTTCTTATTTGGGCGCTGTTAAGTAT 2562
Db 1981 GGAATACATGGCTGTGTGGGCGATCAACACACAGTTTCTTATTTGGGCGCTGTTAAGTAT 2040
Qy 2563 GAGGGGAGAGAGATGAGAGAACTTTTCTATCTGCGCGCTGTCTCAGGGGAGGAACCTG 2622
Db 2041 GAGGGGAGAGAGATGAGAGAACTTTTCTATCTGCGCGCTGTCTCAGGGGAGGAACCTG 2100
Qy 2623 ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCAGCAGCTCTCTGAGTCCCTCCAC 2682
Db 2101 ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCAGCAGCTCTCTGAGTCCCTCCAC 2160
Qy 2683 TGCCTTTACAGACTCGGAACAAAAGTTCCTGACACAAAGTATGGCCCATTTTGAAGAA 2742
Db 2161 TGCCTTTACAGACTCGGAACAAAAGTTCCTGACACAAAGTATGGCCCATTTTGAAGAA 2220
Qy 2743 ATGGGCAATGTGTAGAAAACAGACATGGAGCTTTAGTGTGACATTTTGCATTTAAATTC 2802
Db 2221 ATGGGCAATGTGTAGAAAACAGACATGGAGCTTTAGTGTGACATTTTGCATTTAAATTC 2280

Qy 2803 AGCCGCCACGTGAAGAAAGCTTTCACTGATTGAGGGCAGGACAGACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACGTGAAGAAAGCTTTCACTGATTGAGGGCAGGACAGACAGATCAACATGGAGC 2340
Qy 2863 CCCACCATGGTAGTCTTGTTCAGGTGGGTCCAGCTCAGATGCCATTTGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTAGTCTTGTTCAGGTGGGTCCAGCTCAGATGCCATTTGGCAGATTTCTC 2400
Qy 2923 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGGACCTTAAGTGGAAATCTCGCTG 2982
Db 2401 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGGACCTTAAGTGGAAATCTCGCTG 2460
Qy 2983 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAGACCCCTGAGACGCCCTCTGCTGCTCGAG 3042
Db 2461 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAGACCCCTGAGACGCCCTCTGCTGCTCGAG 2520
Qy 3043 ACCCTGGGTTGGCTGGCTGTGGCTCAGAGTGAAGACTGAAGGACTTGGCTTTGGG 3102
Db 2521 ACCCTGGGTTGGCTGGCTGTGGCTCAGAGTGAAGACTGAAGGACTTGGCTTTGGG 2580
Qy 3103 CTGAGAGCAACACAGACCCCTGACCGAGCTGACCTTGAGCTTCAATGTCTCAGCGATGCT 3162
Db 2581 CTGAGAGCAACACAGACCCCTGACCGAGCTGACCTTGAGCTTCAATGTCTCAGCGATGCT 2640
Qy 3163 GGAGCCAAACACCTTTTCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG 3222
Db 2641 GGAGCCAAACACCTTTTCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG 2700
Qy 3223 CTGTCAGCTGTGGCTCAGCTGTGACTGTGCTGCCAGGACCTGGGCTCTGTGTTAGTGCC 3282
Db 2701 CTGTCAGCTGTGGCTCAGCTGTGACTGTGCTGCCAGGACCTGGGCTCTGTGTTAGTGCC 2760
Qy 3283 AGCCCGACCTGAAGGAGCTAGACCTGACAGAGAAACCTGGATGAGCTTGGCGTGGGA 3342
Db 2761 AGCCCGACCTGAAGGAGCTAGACCTGACAGAGAAACCTGGATGAGCTTGGCGTGGGA 2820
Qy 3343 CTGCTCTGTGAGGGGCTCAGGCTATCTGCTCTGCAAACTCATACGCTTGGGGCTGGACCAG 3402
Db 2821 CTGCTCTGTGAGGGGCTCAGGCTATCTGCTCTGCAAACTCATACGCTTGG----- 2869
Qy 3403 ACAACTCTGATGATGAGATGAGGAGGAACTGAGGGGCCCTGGAGCAGAGAAACCTCAG 3462
Db 2870 ----- 2869
Qy 3463 CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATAGC 3522
Db 2870 -----GGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATAGC 2910
Qy 3523 GGAGAGATGAGTAATAGCACATCTCTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
Db 2911 GGAGAGATGAGTAATAGCACATCTCTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 2970
Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTAAACTCTTGGACGTCGAGCAAGATCTTCCCAATT 3642
Db 2971 GCTTCCCATGTTGCTCAGGCTAATCTAAACTCTTGGACGTCGAGCAAGATCTTCCCAATT 3030
Qy 3643 GCTGAGATGACAGAGAAAGCTCCCGAGGCTAGTACCGGTGGAACTCTTGTGGCTGCT 3702
Db 3031 GCTGAGATGACAGAGAAAGCTCCCGAGGCTAGTACCGGTGGAACTCTTGTGGCTGCT 3090
Qy 3703 TCTCTGCTCTCAAGGGGACCTGATACGAAGCTTTGGGACCTGACGATGACTTCTGG 3762
Db 3091 TCTCTGCTCTCAAGGGGACCTGATACGAAGCTTTGGGACCTGACGATGACTTCTGG 3150
Qy 3763 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGAAGAAAGAACTTGTACCGAGTT 3822
Db 3151 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGAAGAAAGAACTTGTACCGAGTT 3210
Qy 3823 CACTTCCCTGATGCTGGCTCTACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGA 3882
Db 3211 CACTTCCCTGATGCTGGCTCTACCGCTGGCCCAACAGGGTCTCTGCTTGTGATGAGA 3270

Qy	3883	GAAGCGGTGACCGTTCAGATTGAATCTGTGTGGACAGTTCCTGGTGAGATCAAC	3942
Db	3271	GAAGCGGTGACCGTTCAGATTGAATCTGTGTGGACAGTTCCTGGTGAGATCAAC	3330
Qy	3943	CCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGACATCAAGCTGAGCTGGAGCT	4002
Db	3331	CCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGACATCAAGCTGAGCTGGAGCT	3390
Qy	4003	GTGGAAGCTGTGCACCTCCCTCACTTGTGGCTCTCAAGGGGGCCATGTGACACATCC	4062
Db	3391	GTGGAAGCTGTGCACCTCCCTCACTTGTGGCTCTCAAGGGGGCCATGTGACACATCC	3450
Qy	4063	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAAAGCAGCCAGGGTG	4122
Db	3451	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAAAGCAGCCAGGGTG	3510
Qy	4123	GAGCTGCATCACAATGTTCTGGAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAA	4182
Db	3511	GAGCTGCATCACAATGTTCTGGAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAA	3570
Qy	4183	ATGATCCATAATGCCCTGCGCTTCACTTCCCGTCACCTCTGTGTGTGTTTACACCCG	4242
Db	3571	ATGATCCATAATGCCCTGCGCTTCACTTCCCGTCACCTCTGTGTGTGTTTACACCCG	3630
Qy	4243	GTCCATCTGAGGAAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTCCATTGG	4302
Db	3631	GTCCATCTGAGGAAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTCCATTGG	3690
Qy	4303	AGGAATCTGGAGCTCTGCTATGAAAGCCCTGAGAAAGACAGCTGTTCTGGAGTCTAC	4362
Db	3691	AGGAATCTGGAGCTCTGCTATGAAAGCCCTGAGAAAGACAGCTGTTCTGGAGTCTAC	3750
Qy	4363	GTGGCCACTTGGGATCAGGATCAGGCTGCAAGTCAAGACAGAAAGATGAGACTCTG	4422
Db	3751	GTGGCCACTTGGGATCAGGATCAGGCTGCAAGTCAAGACAGAAAGATGAGACTCTG	3810
Qy	4423	GTGTGGAGGCTTGTGTAACCCAGAGATCTCATGCTGCAACTACTCTGATCCCTCA	4482
Db	3811	GTGTGGAGGCTTGTGTAACCCAGAGATCTCATGCTGCAACTACTCTGATCCCTCA	3870
Qy	4483	GCCCGATAGCCGTACTTCACTCTGATGCCCCAGTGTGTCACCTTGTGACACAG	4542
Db	3871	GCCCGATAGCCGTACTTCACTCTGATGCCCCAGTGTGTCACCTTGTGACACAG	3930
Qy	4543	TATCGAGAGCAGCTGATAGCCGAGTGACATCGTGGAGTGTCTTGGACAACTGCAT	4602
Db	3931	TATCGAGAGCAGCTGATAGCCGAGTGACATCGTGGAGTGTCTTGGACAACTGCAT	3990
Qy	4603	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAACACGAGGCCAGC	4662
Db	3991	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAACACGAGGCCAGC	4050
Qy	4663	CAGATCGGAAGCTGTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAGAAATGAGCTC	4722
Db	4051	CAGATCGGAAGCTGTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAGAAATGAGCTC	4110
Qy	4723	TACCAAGCCCTGAAGAGAGCCATCTCCTACCTATTATGAACTCTGGGAGAGGCGAGC	4782
Db	4111	TACCAAGCCCTGAAGAGAGCCATCTCCTACCTATTATGAACTCTGGGAGAGGCGAGC	4170
Qy	4783	AAAAAGGAGCTCTCTGCACTCAGCAGCTGA	4812
Db	4171	AAAAAGGAGCTCTCTGCACTCAGCAGCTGA	4200

RESULT 11
AX089763
LOCUS
DEFINITION
Sequence 3 from Patent WO0116170.
AX089763
ACCESSION
AX089763.1 GI:13443935
KEYWORDS
Homo sapiens (human)

ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 Reed, J.C.									
TITLE	Card proteins involved in cell death regulation									
JOURNAL	Patent: WO 0116170-A 3 08-MAR-2001;									
FEATURES	The Burnham Institute (US)									
source	Location/Qualifiers									
	1. .4194									
	/organism="Homo sapiens"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:9606"									
CDS	1. .4194									
	/note="unnamed protein product"									
	/codon_start=1									
	/protein_id="CAC34900.1"									
	/db_xref="GI:13443936"									
	/translation="MAGGAWRLACYLEFLKKEELKEFOLLANKAHSSSSGETPAQ									
	PEKTSMEVASYLVQAQGEQAWDLALHTWEQMLRSLCAQAGAGHSPFPYSPSE									
	PHLGSPSOTSTAVLMPWIHELPAQCTGSSRRVLRQLPDTSGRRWRISASLDLYQAL									
	PSSDPHSPQESPNAPTSTAVLGSWSPPOPSLAPRQEPAGTQWPDSTSGIYYTE									
	IREREREKGRPPAAVVGCTPQAHTSLQPHHPWEPVSRESLCSWPMKNEFDNQ									
	KFTQLLLQRPHPRSQDELVKRSWPDVVEENRHLIEIRDLFGPGLDTQBERIVLQG									
	AAGIKSTLAEQVKEAMRGQLYGDGFQHFVYFSCRELAQSKVSLAELIGKDGATP									
	APIRQLISRPERLLFILDGVEPGWLOEPSSELSCHWSQPPADALGSLGKTLIP									
	EASFLITARTTALQNLIPSLBQARWVLEGPSESRKEYFYFYRTDERQAIARFLVK									
	SNKELMALCLVPWVSLACTCLMOQMKRKEKLTSTTTCLHLVLAQALQAQPLGP									
	QIRDLCSLAEGIQWKTFLSPDDLRLKHLGDAIISTFLKMGILQEHPIPLSYFIHL									
	QRDFFAAMSYYLDEKGRKHSNCIIDLEKTLIAYGHLGFGASTRFLGLSLDEG									
	EREMENPHCLRSQGRNLQMWPSLQLQLPHLSLESLCYETRNKTLFTQWMAHFE									
	MGMCVETDELLVCTFCIKFSRHVKQLQIEGRQHRSTWPTMVLFRWVPTDAYWQ									
	ILFVLKVTNRLKELDSLNSLSHSAVSKLCTLRPRCLLETLRLACGLTAEDCKDL									
	LAFLGRANOTLTDLSENVLTDAKAKHLQRLRQPSCKLQRLVSCGLTSDCCODL									
	ASVLSASPLKELDLQNNLDDVGRLLCEGRHPACKLIRLKPSWMTPTSGLDTGEM									
	SNSTSSLRQKRSERAAASHVAQNLKLLDVSKIFPTAIEAESSEPVVPVELLCVPS									
	PASQDHLTKPLGTDGDFGPTGVATEWVDKRNLYRVHFPVAGSVYRWPTGLCFVM									
	REAVTVEIEFCVQWDFCEINPQHSVMVAGPLLDIKAPPGVAEAVHLPHFVALQGHV									
	DTSLFQMAHFKKEGMLEKPAVELHVLLENPSFPLGLVKMLHNAIRLPIPTSVV									
	LLYHRVHPEEVTFLYLIIPSDCSIRELELCYRSPGEOOLFSEFYVGHLSGIRLQVKD									
	KDETLVWEALVKFGDMLPATLIPPARIAVPSPLDAPQLLHFDVQDREQLIARVTSV									
	EWVLDKLHGHSQEQYERVLAEINTRPQMRKLFSLQSQSWDRKCKDGLYQALKETHPH									
	LIMELWEKSGSKGLPLSS"									
BASE COUNT	984	a	1167	c	1176	g	867	t		
ORIGIN										
Query Match	74.9%; Score 4078; DB 6; Length 4194;									
Best Local Similarity	97.8%; Pred. No. 0;									
Matches 4194; Conservative	0; Mismatches 0; Indels 96; Gaps 2;									
QY	523	ATGGCTGGCGAGCCTGGGCGCGCTGGCCTGTGTTACTTTGGAGTTCTCTGAAGAAGAGGAG	582							
Db	1	ATGGCTGGCGAGCCTGGGCGCGCTGGCCTGTGTTACTTTGGAGTTCTCTGAAGAAGAGGAG	60							
QY	583	CTGAAGAGAGTTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	642							
Db	61	CTGAAGAGAGTTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	120							
QY	643	ACACCCGCTCAGCCAG	702							
Db	121	ACACCCGCTCAGCCAG	180							
QY	703	TATGGGAG	762							
Db	181	TATGGGAG	240							
QY	763	TCACCTGTGCGCCCAAGCCAG	822							
Db	241	TCACCTGTGCGCCCAAGCCAG	300							
QY	823	AGTGAACCCCACTGGGGTCTCCCGAGCAACCCACCTCCACCGAGTGTCTAATGCCCTGG	882							
Db	301	AGTGAACCCCACTGGGGTCTCCCGAGCAACCCACCTCCACCGAGTGTCTAATGCCCTGG	360							

QY 883 ATCCATGAATTCGCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG 942
Db 361 ATCCATGAATTCGCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG 420
QY 943 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTT 1002
Db 421 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCGAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCATCCACA 1062
Db 481 CCAAGCTCCCGAGACCATGAGTCTCAAGCCAGGAGTCAACCAAGCCGCCCATCCACA 540
QY 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCACCCAGAGAGAGGAG 1122
Db 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGGAGTCAACCAAGCCGCCCATCCACA 600
QY 1123 GCTCTGGACCAATGGCTCTGATGAAGCTCAGGAATTTACTACACAGAAATCAGA 1182
Db 601 GCTCTGGACCAATGGCTCTGATGAAGCTCAGGAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGCAGGCCGCCCATGGCAGCGGTGTAGGAAGC 1242
Db 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGCAGGCCGCCCATGGCAGCGGTGTAGGAAGC 720
QY 1243 CCCCCACAGGCGCACACCGCTTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGA 1302
Db 721 CCCCCACAGGCGCACACCGCTTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACAGAG 1362
Db 781 GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACAGAG 840
QY 1363 CTGCTACTTCTACAAAGACTCACCCAGAGCCAGAGTCCCTGTCAAGAGAGAGCTGG 1422
Db 841 CTGCTACTTCTACAAAGACTCACCCAGAGCCAGAGTCCCTGTCAAGAGAGAGCTGG 900
QY 1423 CTGATTAATGTGAGAGAGAAATCAGAGACATTTAAATTGAGATCAGAGACTTTATTTGGCCCA 1482
Db 901 CTGATTAATGTGAGAGAGAAATCAGAGACATTTAAATTGAGATCAGAGACTTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACTGCTGATGATCATACTCAGGGGCTGTCTGGAATTTGGGAAG 1542
Db 961 GGCCTGGATACCAAGAACTGCTGATGATCATACTCAGGGGCTGTCTGGAATTTGGGAAG 1020
QY 1543 TCAACACTGSCAGGAGTGAAGGAAGCTGGGGAGAGGCCAGCTGATGGGACCGC 1602
Db 1021 TCAACACTGSCAGGAGTGAAGGAAGCTGGGGAGAGGCCAGCTGATGGGACCGC 1080
QY 1603 TTCCAGCATGTCTTTACTTTCACTGTCAGAGAGCTGGCCCACTCAGAGTGGTGTGCTC 1662
Db 1081 TTCCAGCATGTCTTTACTTTCACTGTCAGAGAGCTGGCCCACTCAGAGTGGTGTGCTC 1140
QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGCGCTCCCATTTAGACAGATCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCGCGCTCCCATTTAGACAGATCTGTCT 1200
QY 1723 AGSCCAGAGCGGCTCTTCTCATCTCGATGTTAGATGAGCCAGGATGGGCTGTGAG 1782
Db 1201 AGSCCAGAGCGGCTCTTCTCATCTCGATGTTAGATGAGCCAGGATGGGCTGTGAG 1260
QY 1783 GAGCCAGATTTCTGAGCTCTGCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGTGCTG 1842
Db 1261 GAGCCAGATTTCTGAGCTCTGCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGTGCTG 1320
QY 1843 GGCAGTTTCTGGGAAAAACTATATTTCCGAGGCATCTCTCTGATCAGGCTCGGACC 1902
Db 1321 GGCAGTTTCTGGGAAAAACTATATTTCCGAGGCATCTCTCTGATCAGGCTCGGACC 1380
QY 1903 ACAGCTCTGAGNAACTTCTTCTTTGGAGCAGGACGTTGGGTAGAGGTCTGGGG 1962
Db 1381 ACAGCTCTGAGNAACTTCTTCTTCTTTGGAGCAGGACGTTGGGTAGAGGTCTGGGG 1440

QY 1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTCA CAGATGAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTCA CAGATGAAGGCAAGCA 1500
QY 2023 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
Db 1501 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560
QY 2083 TGGGTGTCTGGCTGGCCTGCACTTGGCTGATGTCAGCAGATGAAGCGGAAGGAAAACTC 2142
Db 1561 TGGGTGTCTGGCTGGCCTGCACTTGGCTGATGTCAGCAGATGAAGCGGAAGGAAAACTC 1620
QY 2143 ACACGTACTTCCAAGACACACCAACCTCTGTCTACATTTACCTTGGCCAGCTCTCNA 2202
Db 1621 ACACGTACTTCCAAGACACACCAACCTCTGTCTACATTTACCTTGGCCAGCTCTCNA 1680
QY 2203 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGCTGAGGCACTGG 2262
Db 1681 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGCTGAGGCACTGG 1740
QY 2263 CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGCCATC 2322
Db 1741 CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGCCATC 1800
QY 2323 ATCTCCACTTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
Db 1801 ATCTCCACTTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860
QY 2383 TTCATTACCTCTGTTTCCAGAGTTCCTTTCAGCAATGTCTCTGAGGATGAG 2442
Db 1861 TTCATTACCTCTGTTTCCAGAGTTCCTTTCAGCAATGTCTCTGAGGATGAG 1920
QY 2443 AAGGGGAGAGGTAAACATTTCTAAATTCATAGATTTGGAAAAAGACGCTAGAAGCATAT 2502
Db 1921 AAGGGGAGAGGTAAACATTTCTAAATTCATAGATTTGGAAAAAGACGCTAGAAGCATAT 1980
QY 2503 GGAATACATGGCCTGTTTGGGGCATCAACCAACGTTTCTTCTTGGGCTGTAAAGTAT 2562
Db 1981 GGAATACATGGCCTGTTTGGGGCATCAACCAACGTTTCTTCTTGGGCTGTAAAGTAT 2040
QY 2563 GAGGGGAGAGAGATGGAGAACATTTCTACTCCGGCTGTCTCAGGGGAGGACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAACATTTCTACTCCGGCTGTCTCAGGGGAGGACCTG 2100
QY 2623 ATGCACTGGTCCCTCCCTGCACTGTCTGTCAGCCACACACTCTCTGAGTCCCTCCAC 2682
Db 2101 ATGCACTGGTCCCTCCCTGCACTGTCTGTCAGCCACACACTCTCTGAGTCCCTCCAC 2160
QY 2683 TGCTTGTACGAGACTCGGAAACAAACGTTTCTGACACAAGTATGGCCCATTTTCGAAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAAACAAACGTTTCTGACACAAGTATGGCCCATTTTCGAAGAA 2220
QY 2743 ATGGGCATGTGTAGAAACACAGATGGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2802
Db 2221 ATGGGCATGTGTAGAAACACAGATGGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2280
QY 2803 AGCCGCCACGTCGAAGAGCTTCAGCTGATTTGAGGCGAGCAGACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACGTCGAAGAGCTTCAGCTGATTTGAGGCGAGCAGCAGATCAACATGGAGC 2340
QY 2863 CCCACCATGGTATCTCTTTCAGTGGGTCCAGATCAGAGATGCCTATTGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTATCTCTTTCAGTGGGTCCAGATCAGAGATGCCTATTGGCAGATTTCTC 2400
QY 2923 TTCTCCGTCTCAAGGTTCACAGAAACCTGAAGAGCTGGACCTAAGTGGAAACCTGCTG 2982
Db 2401 TTCTCCGTCTCAAGGTTCACAGAAACCTGAAGAGCTGGACCTAAGTGGAAACCTGCTG 2460
QY 2983 AGCCACTCTGCAAGTCAAGAGCTTTTGTAAAGCCCTGAGACGCCCTCGCTCCCTGGAG 3042
Db 2461 AGCCACTCTGCAAGTCAAGAGCTTTTGTAAAGCCCTGAGACGCCCTCGCTCCCTGGAG 2520
QY 3043 ACCCTCGGTTGGCTGTGGCTCTCAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 3102

2521	Db		ACCCTGGGTTGGCTGGCTGGCTGCACAGCTGAGACTGCAAGACCTTGCCCTTTGGG	2581
3103	Qy		CTGAGAGCAACACAGACCTTGACCGAGCTGAGACTGAGCTTCAATGTCTCACGAGTGCT	3162
2581	Db		CTGAGAGCCAACACAGACCTTGACCGAGCTGAGACTGAGCTTCAATGTCTCACGAGTGCT	2640
3163	Qy		GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCAG	3222
2641	Db		GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCAG	2700
3223	Qy		CTGTCTAGCTGTGGCCCTCACTCTGACTGTGCCAGAGACCTTGGCTCTGTGCTTAGTGCC	3282
2701	Db		CTGTCTAGCTGTGGCCCTCACTCTGACTGTGCCAGAGACCTTGGCTCTGTGCTTAGTGCC	2760
3283	Qy		AGCCCCAGCCTGAAGGAGCTAGACCTCCACAGACAAACCTGGATGACGTTGGCGTCGA	3342
2761	Db		AGCCCCAGCCTGAAGGAGCTAGACCTCCACAGACAAACCTGGATGACGTTGGCGTCGA	2820
3343	Qy		CTGTCTGTGAGGGGCTCAGGCATCTCTGCCCTGCAAACTCATACGCCCTGGGCTGGACCAG	3402
2821	Db		CTGTCTGTGAGGGGCTCAGGCATCTCTGCCCTGCAAACTCATACGCCCT-----	2867
3403	Qy		ACAACTCTGAGTGATGATGAGCGCAGAACTGAGGGCCCTGGAGCAGGAGAAACCTTCAG	3462
2868	Db		-----	2867
3463	Qy		CTGCTCATCTTCAGCAGACCGAAACCAAGTGTGATGATGACCCCTACTGAGGGCCCTGGATACG	3522
2868	Db		-----GAAACCAAGTGTGATGATGACCCCTACTGAGGGCCCTGGATACG	2907
3523	Qy		GGAGAGTAGTAAATAGACATCTCACTCAAGCGGCAGAGACTCGCATCAGAGAGGGCG	3582
2908	Db		GGAGAGTAGTAAATAGACATCTCACTCAAGCGGCAGAGACTCGCATCAGAGAGGGCG	2967
3583	Qy		GCTTCCCATGTTGCTCAGGCTTAATCTCAAACTCTCTGGAACGTGAGCAAGATCTTCCCAATT	3642
2968	Db		GCTTCCCATGTTGCTCAGGCTTAATCTCAAACTCTCTGGAACGTGAGCAAGATCTTCCCAATT	3027
3643	Qy		GCTGAGATTCGACAGAGAAAGCTCCACAGAGGTAGTACCGGTGGAACTCTTTGGCGTGCCCT	3702
3028	Db		GCTGAGATTCGACAGAGAAAGCTCCACAGAGGTAGTACCGGTGGAACTCTTTGGCGTGCCCT	3087
3703	Qy		TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGACTGACGATGACTCTCTGG	3762
3088	Db		TCTCCTGCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGACTGACGATGACTCTCTGG	3147
3763	Qy		GGCCCCACGGGCGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3822
3148	Db		GGCCCCACGGGCGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3207
3823	Qy		CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
3208	Db		CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3267
3883	Qy		GAAGCGGTGACCGTTGAGTTGAATCTGTGTGGGACACAGTTCTCTGGGTGAGATCAAC	3942
3268	Db		GAAGCGGTGACCGTTGAGTTGAATCTGTGTGGGACACAGTTCTCTGGGTGAGATCAAC	3327
3943	Qy		CCACAGCACAGCTGGATGGTGACAGGGCCCTCTGCTGGACATCAAGGCTGAGCCCTGGAGCT	4002
3328	Db		CCACAGCACAGCTGGATGGTGACAGGGCCCTCTGCTGGACATCAAGGCTGAGCCCTGGAGCT	3387
4003	Qy		GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCC	4062
3388	Db		GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCC	3447
4063	Qy		CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAGACGACCGAGGCTG	4122
3448	Db		CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAGACGACCGAGGCTG	3507
4123	Qy		GAGCTGCATCACATAGTCTCTGGAAACCCCGACTTCTCCCTTTGGGAGTCTCTCTGAAA	4182

3508	3508	Db	GAGCTGCATCACATAGTCTTGGAAAAACCCAGGCTTCTCCCGCTTGGAGTCTCTCTGAAA	3556
4183	4183	Qy	ATGATCATCAATATGCCCTCGCGCTTCATTCCCGTCACTCTGTGTGTGTGCTTTTACCAACCGC	4242
3568	3568	Db	ATGATCATCAATATGCCCTCGCGCTTCATTCCCGTCACTCTGTGTGTGTGCTTTTACCAACCGC	3627
4243	4243	Qy	GTCCATCTCTGAGGAAGTCACTTTCACCTCTACCTGATCCCAAGTGAAGTCTGCTCCATTCGG	4302
3628	3628	Db	GTCCATCTCTGAGGAAGTCACTTTCACCTCTACCTGATCCCAAGTGAAGTCTGCTCCATTCG-	3686
4303	4303	Qy	AAGAACTCGAGGCTCTGCTATCGAAGCCCTGGAGAGACACAGCTGTTCTCGAGAGTCTTAC	4362
3687	3687	Db	--GGAATCTGAGGCTCTGCTATCGAAGCCCTGGAGAGACACAGCTGTTCTCGAGAGTCTTAC	3744
4363	4363	Qy	GTTCGCCACTTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAGAAAGATGAGACTCTG	4422
3745	3745	Db	GTTCGCCACTTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAGAAAGATGAGACTCTG	3804
4423	4423	Qy	GTGTGGAGAGCCCTGTGTGAACACAGAGAGATCTCATGCTGCAACTACTCTGTGATCCCTCCA	4482
3805	3805	Db	GTGTGGAGAGCCCTGTGTGAACACAGAGAGATCTCATGCTGCAACTACTCTGTGATCCCTCCA	3864
4483	4483	Qy	GCCCGCATAGCCGTACCTTCACCTCTGGATGCCCGCCAGTGTGTCACCTTTGTGGACCCAG	4542
3865	3865	Db	GCCCGCATAGCCGTACCTTCACCTCTGGATGCCCGCCAGTGTGTCACCTTTGTGGACCCAG	3924
4543	4543	Qy	TATCGAGAGAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTCTGGACAACACTGCAT	4602
3925	3925	Db	TATCGAGAGAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTCTGGACAACACTGCAT	3984
4603	4603	Qy	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTGGCTGAGAACACAGAGGCCACGC	4662
3985	3985	Db	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTGGCTGAGAACACAGAGGCCACGC	4044
4663	4663	Qy	CAGATCGGAAGCTGTTCAGCTTTAGCCAGTCTCTGGACCGGAAGTGCACAAGATGACATC	4722
4045	4045	Db	CAGATCGGAAGCTGTTCAGCTTTAGCCAGTCTCTGGACCGGAAGTGCACAAGATGACATC	4104
4723	4723	Qy	TACCAAGCCCTGAAGGAGACCCATCTTCACTCATTTATGGAACTCTGGGAGAAGGGCAGC	4782
4105	4105	Db	TACCAAGCCCTGAAGGAGACCCATCTTCACTCATTTATGGAACTCTGGGAGAAGGGCAGC	4164
4783	4783	Qy	AAAAAGGAGTCTCTGCCACTCAGAGCTGA	4812
4165	4165	Db	AAAAAGGAGTCTCTGCCACTCAGAGCTGA	4194

RESULT 12
AF229062

AF229062	AF229062	4332 bp	linear	PRI 02-FEB-2001
LOCUS				
DEFINITION	Homo sapiens NAC-delta splice variant (NAC) mRNA, complete cds,			
	alternatively spliced.			
ACCESSION	AF229062			
VERSION	AF229062.1	GI:12856110		

KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens

ORGANISM
Homo sapiens
Metazoa; Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 4332)
AUTHORS
Chu, Z.-L., Pio, F., Xie, Z., Godzik, A. and Reed, J. C.

TITLE NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c pathway for apoptosis

JOURNAL
REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 4332)
Chu, Z.-L. and Reed, J.C.

CHU, Z.-L. and REED, G.C.
 TITLE Direct Submission
 JOURNAL Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The
 Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA
 92037 USA

FEATURES	Location/Qualifiers
source	1 4332
92037, USA	

GAGCTGCATCACATAGTCTCTGGAAAACCCACAGCTCTCTCCCCCTTTGGGAGTCTCTCTGAAA 4182

gene	mol_type="mRNA"	Db	361	ATCCATGAATTGCCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTGACAGCTG	420
CDS	/db_xref="taxon:9606"	Qy	943	CCTGACACATCTGCAGCCCGCTGGAGAGAAATCTTGCCTCACTCTCTACCAAGCTCTT	1002
	/chromosome="11"	Db	421	CCTGACACATCTGCAGCCCGCTGGAGAGAAATCTTGCCTCACTCTCTACCAAGCTCTT	480
CDS	/cell_type="activated Jurkat T cells"	Qy	1003	CCAGCTCCCCAGACCAATGAGTCTCCAAAGCCAGGAGTCAACCAAGCCCCCAATCCACA	1062
	1..4332	Db	481	CCAAAGTCCCCAGACCAATGAGTCTCCAAAGCCAGGAGTCAACCAAGCCCCCAATCCACA	540
CDS	/gene="NAC"	Qy	1063	GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCAGCCCTAGCACCCAGAGCAGGAG	1122
	1..4332	Db	541	GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCAGCCCTAGCACCCAGAGCAGGAG	600
CDS	/codon_start=1	Qy	1123	GCTCTCGGACCCCAATGGCTCTCGATGAAACGTGAGAAATTTACTACACAGAAATCAGA	1182
	/product="NAC-delta splice variant"	Db	601	GCTCTCGGACCCCAATGGCTCTCGATGAAACGTGAGAAATTTACTACACAGAAATCAGA	660
CDS	/protein_id="AAK00751.1"	Qy	1183	GAAAGAGAGAGAGAAATCAGAAAGGAGGCCCCCATGGGAGCGGTGGTAGGAACG	1242
	/db_xref="GI:12656111"	Db	661	GAAAGAGAGAGAGAAATCAGAAAGGAGGCCCCCATGGGAGCGGTGGTAGGAACG	720
CDS	/translation="MAGGAWGLRLCYLEFLKKEBELFQLLLANKAHSSSSSGETPAQ	Qy	1243	CCCCACAGGCGCACACGAGCCTACAGCCCAACACACCCATCGGGAGCCTTCTGTGAGA	1302
	PEKTSMEVASLVQAQGEORAWDLALHTWEOMGIRSLCAQAOEGAGHSPSPSPSE	Db	721	CCCCACAGGCGCACACGAGCCTACAGCCCAACACACCCATCGGGAGCCTTCTGTGAGA	780
CDS	PHLGSPQPTAVLMPWIHELPAQCTQGSERRVLRQLPDTSGRRWRREISASLLIYQAL	Qy	1303	GAGAGCCTCTGTTCCACATGSCCTCGAAATAATGAGGATTTAAACCAAAAATTTACACAG	1362
	FSPSPSPQSPQNPATSTAVLGSWSPPQSLAPREQAPGTQPLDSTGSIYYTE	Db	781	GAGAGCCTCTGTTCCACATGSCCTCGAAATAATGAGGATTTAAACCAAAAATTTACACAG	840
CDS	TREREXSEKSPAAVSTVPQAHTSLQHPHPWESPRESLCTSPWKNEDNQ	Qy	1363	CTGCTACTTTCACAAAGACCTCACCCAGAAAGCCAGATCCCCCTGGTCAAGAGAAAGCTG	1422
	KATQILLQRPHRSQDLVKRSPVVEENRHLEIRDLFGDLDQEPRIVLQ	Db	841	CTGCTACTTTCACAAAGACCTCACCCAGAAAGCCAGATCCCCCTGGTCAAGAGAAAGCTG	900
CDS	FFQGLLLQARVKEAMRGQLYGRFOHVPYFSCRELAQKVVSIAELIKGDTATP	Qy	1423	CCTGATATGTGGAGGAATTCGAGGACATTTAATTGAGATCAGAGACTTTATTGGCCCA	1482
	APIROILSRPELLFLLQVDEPQWVLPQSPSELCLHWSQPOPADALLGSLGKTLIP	Db	901	CCTGATATGTGGAGGAATTCGAGGACATTTAATTGAGATCAGAGACTTTATTGGCCCA	960
CDS	ENSEFLIARTTALNLPISLEQAQRWVEVLGSESSRKEYFYRTPIERQAIRAFRLVK	Qy	1483	GSCCTGGAATACCCAAAGAACCTCGCATAGTCTATCTGACGAGGGCTGCTGGAATTGGGAAG	1542
	SNKELIATLVPWVSWLACTCLMQMRKKEKTLTSTKTTTTLCHLYLAQALQAQPLQ	Db	961	GSCCTGGAATACCCAAAGAACCTCGCATAGTCTATCTGACGAGGGCTGCTGGAATTGGGAAG	1020
CDS	QFQDFCSAABGIRWQKTLFSPDLRKHLDGAIISTFKMGILQHEHPLPSYFSTHL	Qy	1543	TCAACACATCGGCAGCGAGTGAAGAAAGCCCTGGGGAGAGCCAGCTGTATGGGGACCGC	1602
	CLPDEFAAMSVLDEKGRGHSNCIIDLEKTLIYAGIHGLFGASTFRLLGLSDG	Db	1021	TCAACACATCGGCAGCGAGTGAAGAAAGCCCTGGGGAGAGCCAGCTGTATGGGGACCGC	1080
CDS	MEMENI FCHRLSQGRNLQWVPSLQLLLOPHLSLESLHCLVETRNKTFELQVMAHP	Qy	1603	TTCCAGCATGTCTTCTACTTCAGCTGCAGAGAGCTGCCCCCAGTCCAAAGTGGTGAAGTCTC	1662
	MGMCVETMELLVCTFCIFKSRHVKKLQILEGROHRSWTSPVWVFRVPTVDTAQ	Db	1081	TTCCAGCATGTCTTCTACTTCAGCTGCAGAGAGCTGCCCCCAGTCCAAAGTGGTGAAGTCTC	1140
CDS	TLFVSVKTRNLKELDLSGNSLSHSAVKSICKTLRRRCLELRLRAGCLTAEBCND	Qy	1663	GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTCTGCT	1722
	LAFGRANQITLDELSDNLVDAGKHLQRLQRPQSKLQRLQVSCGLTSDCCQDL	Db	1141	GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTCTGCT	1200
CDS	ASVLSASLKLKELDILQNNLDDVGRVLLCEGLRHPACKLRILGKPSVMTPEGLDTE	Qy	1723	AGGCCAGAGCGGCTGCTTTCATCTCGATGGTGTAGATGAGCCAGATGGGTCTTTCAG	1782
	MSNTSSSLQRLQAGSRAASHVAQNLKLDVSKIFPIABIABESSPEVVPVVELLCVP	Db	1201	AGGCCAGAGCGGCTGCTTTCATCTCGATGGTGTAGATGAGCCAGATGGGTCTTTCAG	1260
CDS	SPASQDLHTPLGTDDEFWGPTEVVDVVEKKNLYRVHFPVAGSVRPWPTGLCFV	Qy	1783	GAGCCGAGTTCGTGAGCTGTGTGCACTGAGGCCAGCCACAGCCGCGGATGCACTGCTG	1842
	WREAVTVLEPCVMDQFGEINPQHSWMVAGPLDIIKAEPCVNAEHLPHFVALQGH	Db	1261	GAGCCGAGTTCGTGAGCTGTGTGCACTGAGGCCAGCCACAGCCGCGGATGCACTGCTG	1320
CDS	VDTSLFQWAFKEBGLLEKPARVELHHIVLENSFSPGLLVKLMHNAURLFIPVTSV	Qy	1843	GSCAGTTCGTGGGAAACCTATACTTCCCGAGGCATCTTCTCTGATCAGCGCTCGGACC	1902
	VLLXHVHPEEVTFLYLIIPSDCSIRKAIIDLENMKFQFVRHKPPPLPLYMGCRVTY	Db	1321	GSCAGTTCGTGGGAAACCTATACTTCCCGAGGCATCTTCTCTGATCAGCGCTCGGACC	1380
CDS	SGSGSMLEIILPKELELCYRSGEDQLFSEFYVHGLSGIRLOVKQKDETLVWEALV	Qy	1903	ACAGCTCTGCAGAAACCTCATTTCTTTTGGAGCAGGCACTTTGGGTAGAGGTCTTGGGG	1962
	KPGDLMPATTLIIPPARIAVPSPLDAPOLLHFVDQYRQLIARVTSVEVLDLKHGOVL	Db	1381	ACAGCTCTGCAGAACTCATTTCTTTTGGAGCAGGCACTTTGGGTAGAGGTCTTGGGG	1440
CDS	SOEOYERVLAEENTRPSQMRKLFSLSQSWDRCKDGLYQALKEPHPLIMELWEKSKK	Qy	1963	TTCTCTGAGTCCAGCAGAAAGAAATTTTCTTACAGATATTTTTCAGATGAAAGGCAAGCA	2022
	GLLPLSS"	Db			

Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGCAAGCA 1500
 Qy 2023 ATTAGAGCCTTTAGGTGTGTCAAATCAAAAGAGCTCTGGGCGCTGTGTGTGGCC 2082
 Db 1501 ATTAGAGCCTTTAGGTGTGTCAAATCAAAAGAGCTCTGGGCGCTGTGTGTGGCC 1560
 Qy 2083 TGGGTCTCTGGCTGGCCCTGCATCTTGGCTGATGACAGATGAAGCGGAAGAAACATC 2142
 Db 1561 TGGGTCTCTGGCTGGCCCTGCATCTTGGCTGATGACAGATGAAGCGGAAGAAACATC 1620
 Qy 2143 AACTGACTTCAAGACACCAACACCTCTGTCTATCATTTACTTGGCCCGAGCTCTCCAA 2202
 Db 1621 AACTGACTTCAAGACACCAACACCTCTGTCTATCATTTACTTGGCCCGAGCTCTCCAA 1680
 Qy 2203 GCTCAGCCTATGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGGCATCTGG 2262
 Db 1681 GCTCAGCCTATGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGGCATCTGG 1740
 Qy 2263 CAAAAAAGACCCCTTTTCAAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC 2322
 Db 1741 CAAAAAAGACCCCTTTTCAAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGCGCATC 1800
 Qy 2323 ATCTCACCTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCTCTGAGCTACAGC 2382
 Db 1801 ATCTCACCTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCTCTGAGCTACAGC 1860
 Qy 2383 TTCAATCACTCTGTTTCCAGAGTCTTTTGCAGCAATCTCTATGTCTTGGAGATGAG 2442
 Db 1861 TTCAATCACTCTGTTTCCAGAGTCTTTTGCAGCAATCTCTATGTCTTGGAGATGAG 1920
 Qy 2443 AAGGGGAGAGGTAAACATTTCTAATTCATCATAGATTTGGAAAGACGCTAGAAGCATAT 2502
 Db 1921 AAGGGGAGAGGTAAACATTTCTAATTCATCATAGATTTGGAAAGACGCTAGAAGCATAT 1980
 Qy 2503 GGAATACATGGCTGTTTGGGCGATCAACACACGTTTCTATTTGGGCTGTATTAAGTAT 2562
 Db 1981 GGAATACATGGCTGTTTGGGCGATCAACACACGTTTCTATTTGGGCTGTATTAAGTAT 2040
 Qy 2563 GAGGGGAGAGAGATGAGACATCTTTCACTGCGGCTGTCTCAGGGGAGGACCTG 2622
 Db 2041 GAGGGGAGAGAGATGAGACATCTTTCACTGCGGCTGTCTCAGGGGAGGACCTG 2100
 Qy 2623 ATGCACTGGTCCCGTCCCTGCACTGCTGTGACGACACATCTCTGAGTCCCTCCAC 2682
 Db 2101 ATGCACTGGTCCCGTCCCTGCACTGCTGTGACGACACATCTCTGAGTCCCTCCAC 2160
 Qy 2683 TGCTTTGACGAGCTCGGAACAAACGTTCTTCTGACACAAAGTATGAGCCCATTTGGAAGAA 2742
 Db 2161 TGCTTTGACGAGCTCGGAACAAACGTTCTTCTGACACAAAGTATGAGCCCATTTGGAAGAA 2220
 Qy 2743 ATGGGCATGTGTAGAAACAGACATGGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2802
 Db 2221 ATGGGCATGTGTAGAAACAGACATGGAGCTTTAGTGTGACATTTCTGCAATTAATTC 2280
 Qy 2803 AGCGGCACGTGAAGAGCTTCAAGTGAATGAGGGCAGGACGACACAGATCAACATGGAGC 2862
 Db 2281 AGCGGCACGTGAAGAGCTTCAAGTGAATGAGGGCAGGACGACACAGATCAACATGGAGC 2340
 Qy 2863 CCCACCATGGTATCTGTTTCAAGTGGGTCCAGTCCAGTCCAGTATTTGGCAGATTTCT 2922
 Db 2341 CCCACCATGGTATCTGTTTCAAGTGGGTCCAGTCCAGTCCAGTATTTGGCAGATTTCT 2400
 Qy 2923 TTCTCCGTCTCAAGTTCACGAAACCTTGAAGAGCTGGACCTAGTGGAACTCGCTG 2982
 Db 2401 TTCTCCGTCTCAAGTTCACGAAACCTTGAAGAGCTGGACCTAGTGGAACTCGCTG 2460
 Qy 2983 AGCCACTCTGCACTGAAGAGTCTTTGTAAGACCTGAGACGCTCTGCTGCTCTGGAG 3042
 Db 2461 AGCCACTCTGCACTGAAGAGTCTTTGTAAGACCTGAGACGCTCTGCTGCTCTGGAG 2520
 Qy 3043 ACCCTGCGTGTGGCTGTGGCTCAGAGTGTGGCTCAGAGTGTGGCTGTGGCTGTGGG 3102
 Db 2521 ACCCTGCGTGTGGCTGTGGCTCAGAGTGTGGCTCAGAGTGTGGCTGTGGCTGTGGG 2580

Qy 3103 CTGAGAGCAACCCAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCACGGATGCT 3162
 Db 2581 CTGAGAGCAACCCAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCACGGATGCT 2640
 Qy 3163 GGAGCCAAACACCTTTGCCAGAGCTGAGACAGCGAGCTGCAAGCTACAGCGACTGCAG 3222
 Db 2641 GGAGCCAAACACCTTTGCCAGAGCTGAGACAGCGAGCTGCAAGCTACAGCGACTGCAG 2700
 Qy 3223 CTGGTCACTGTGGCTCAGCTCTGACTGCTGCCAGGACCTGGCTCTGTGTAGTGCC 3282
 Db 2701 CTGGTCACTGTGGCTCAGCTCTGACTGCTGCCAGGACCTGGCTCTGTGTAGTGCC 2760
 Qy 3283 AGCCCCAGCCTGAAGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTGGCTGGCA 3342
 Db 2761 AGCCCCAGCCTGAAGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTGGCTGGCA 2820
 Qy 3343 CTGCTCTGTGAGGGCTCAGGATCTCTGCTGCAAACTCATACGCTGGGCTGGACAG 3402
 Db 2821 CTGCTCTGTGAGGGCTCAGGATCTCTGCTGCAAACTCATACGCTGG----- 2869
 Qy 3403 ACACTCTGAGTGTAGATGAGGCGAGAACTGAGGGCCCTGGAGCAGGAGAACTCAG 3462
 Db 2870 ----- 2869
 Qy 3463 CTGCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAGC 3522
 Db 2870 -----GGAAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAGC 2910
 Qy 3523 GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCGAGAGACTCGGATCAGAGAGGCG 3582
 Db 2911 GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCGAGAGACTCGGATCAGAGAGGCG 2970
 Qy 3583 GCTTCCCATGTTGCTCAGGCTAACTCTCAACTCTGAGACCTGAGCAAGATCTTCCCAATT 3642
 Db 2971 GCTTCCCATGTTGCTCAGGCTAACTCTCAACTCTGAGACCTGAGCAAGATCTTCCCAATT 3030
 Qy 3643 GCTGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGCTGCT 3702
 Db 3031 GCTGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGCTGCT 3090
 Qy 3703 TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTGGGGACTGACGATCTTCTGG 3762
 Db 3091 TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTGGGGACTGACGATCTTCTGG 3150
 Qy 3763 GGCCCAAGGGGCTGTGGCTACTGAGGTAGTGTGACAAAGAAAGAACTTGTACCGATT 3822
 Db 3151 GGCCCAAGGGGCTGTGGCTACTGAGGTAGTGTGACAAAGAAAGAACTTGTACCGATT 3210
 Qy 3823 CACTTCCCTGTAGCTGGCTCTTACCGCTGGCCCAACCGGCTCTCTGCTTTGTGATGAGA 3882
 Db 3211 CACTTCCCTGTAGCTGGCTCTTACCGCTGGCCCAACCGGCTCTCTGCTTTGTGATGAGA 3270
 Qy 3883 GAAGCGGTACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3942
 Db 3271 GAAGCGGTACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3330
 Qy 3943 CCACAGCAGCTGAGTGGTGGCAGGGCTCTGCTGACATCAAGGCTGAGGCTGGAGCT 4002
 Db 3331 CCACAGCAGCTGAGTGGTGGCAGGGCTCTGCTGACATCAAGGCTGAGGCTGGAGCT 3390
 Qy 4003 GTGGAAGCTGTGCACCTCCCTCACTTTCTGCTGCTCTCCAAAGGGGCGCATGTGGACACATCC 4062
 Db 3391 GTGGAAGCTGTGCACCTCCCTCACTTTTGTGCTCTCCAAAGGGGCGCATGTGGACACATCC 3450
 Qy 4063 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGAGTGTCTCTGGAGAACCCAGCGGTG 4122
 Db 3451 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGAGTGTCTCTGGAGAACCCAGCGGTG 3510
 Qy 4123 GAGCTGCAATCAGATGTTCTGGAACCCAGAGTCTTCTCCCTTGGGAGTCTCTCTGAAA 4182
 Db 3511 GAGCTGCAATCAGATGTTCTGGAACCCAGAGTCTTCTCCCTTGGGAGTCTCTCTGAAA 3570

QY	4183	ATGATCCATAATGCGCTTCATTCCTCCGTCACCTCTGTGGTGTGTTTACCAACCGC	4242
Db	3571	ATGATCCATAATGCGCTTCATTCCTCCGTCACCTCTGTGGTGTGTTTACCAACCGC	3630
QY	4243	GTCCATCTCGAGGAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTCCATTCGG	4302
Db	3631	GTCCATCTCGAGGAGTCACTTCCACCTTACCTGATCCCAAGTGAAGTCTCCATTCGG	3690
QY	4303	-----	4302
Db	3691	AAGGCCATAGATGATCTAGAAATGAATTCAGATTGTGCGAATCCACAAGCCACCCCGC	3750
QY	4303	-----	4302
Db	3751	CTGACCCCACTTATATGCGCTGCTGCTTACACTGTCTGTGGTCTGGTTTACGAGGATGCTG	3810
QY	4303	-----	4302
Db	3811	GAATTAATCTCCCAAGGAATGAGAGCTCTGCTATCGAAGCCCTGGAGAACACAGCTGTTC	3870
QY	4351	TGGAGTCTTACCTTGGCCACTTGGGATCAGGATCAGGATCAGGATGAAAGACAGAA	4410
Db	3871	TGGAGTCTTACCTTGGCCACTTGGGATCAGGATCAGGATGAAAGACAGAA	3930
QY	4411	GATGAGACTCTGGTGTGGGAGGCTTTGGTGAACACGAGGATCTCATGCTGCAACTACT	4470
Db	3931	GATGAGACTCTGGTGTGGGAGGCTTTGGTGAACACGAGGATCTCATGCTGCAACTACT	3990
QY	4471	CTGATCCCTCCAGCCGATAGCCGTACCTTACCTCTGATGATCCCGCCAGTGTGTCAC	4530
Db	3991	CTGATCCCTCCAGCCGATAGCCGTACCTTACCTCTGATGATCCCGCCAGTGTGTCAC	4050
QY	4531	TTTGTGACAGTATGAGAGAGCTGATAGCCGAGTGACATCGGTGAGGTTGTCTTG	4590
Db	4051	TTTGTGACAGTATGAGAGAGCTGATAGCCGAGTGACATCGGTGAGGTTGTCTTG	4110
QY	4591	GACAACTGATGAGAGTGTGTGAGCCAGGAGCAGTACGAGAGGTTGCTGCTGAGAAC	4650
Db	4111	GACAACTGATGAGAGTGTGTGAGCCAGGAGCAGTACGAGAGGTTGCTGCTGAGAAC	4170
QY	4651	ACAGGCCAGCAGATGCGGAAGCTGTTGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC	4710
Db	4171	ACAGGCCAGCAGATGCGGAAGCTGTTGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC	4230
QY	4711	AAAGATGAGTCTTACCAAGCCCTGAGGAGACCCATCTCCTCCTCCTCCTCCTCCTCCT	4770
Db	4231	AAAGATGAGTCTTACCAAGCCCTGAGGAGACCCATCTCCTCCTCCTCCTCCTCCTCCT	4290
QY	4771	GAGAGGGGAGCAAAAGGAGTCTCCTGCCACTCAGCAGCTGA	4812
Db	4291	GAGAGGGGAGCAAAAGGAGTCTCCTGCCACTCAGCAGCTGA	4332
RESULT 13			
LOCUS	AX089765	4329 bp	DNA linear PAT 21-MAR-2001
DEFINITION	Sequence 5 from Patent WO0116170.		
ACCESSION	AX089765		
VERSION	AX089765.1	GI:13443937	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Reed, J.C.		
TITLE	Card proteins involved in cell death regulation		
JOURNAL	Patent: WO 0116170-A 5 08-MAR-2001;		
FEATURES	The Burnham Institute (US)		
source	Location/Qualifiers		
	1. .4329		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		

CDS

/db_xref="taxon:9606"

1. .4329

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC34901.1"

/db_xref="GI:13443938"

/translation="MAGGAWRLACYLEFLKKEELKEFOLLANKAHRSRSSSGETPAQPKTSGMEVASYLVAQYGEQAWDLALHTEQMLGRSLCAQAEAGHSPSPYPSPSEPHLGSPQSTAVLMPWILHELPAAGCTQGSRRVLRQPDTSGRWRRIISALLYQALPSSDPSPSOPSNAPTSTAVLGSWGPPOPSLAAPROEAPGTOWPLDETSGIYYTEIIRERERKSEKRPAAVGTPOAHTSLOPHHPHWPESVRESLCTSPWKNEDFNOKFTQLLLQRPSPQDLVRSWPDYVEENRHLIEIRDLPGLDTEPRLVILQQAAGIKSLARQKEAWGRQYIDRFQHFVTFSCRELAQSKVSWSLAELIKDGTATPAPIRQLISPERLLFILDVDEPPWVQPSSELCHWSQPOFADLLSGSLGKILPEASFILITARTALQNLIPSEARWVEVLFSESSKEYFYFYFDEQAIARFALVKSNKELWALCLVPWSWLACTCMQOMKREKLTITSTKTTTLLCLHLAQAQKPLGPLQLRDLCSLAAGSIWOKKTLFSPDDLRKHGLDGAIIISTFLKMGILIOEHPITLPSYFHLCFQEPFAKMSVLEDEKGRKHSNCIIDEKTEAYGIHGLFGASTPFLGLLSDGGEREMENIFHRLSQGRNLMQWPSLQLLQPHSLHSLHCLYETRNKTFITQWMAHPEEMGMCVETDMEILLVCTFCIKFSRHWKQLQIEGRHRSWSPSTMVFLFRWPVTDAYWQILFSLVLRNLKELDLSNLSHSAKSLKTLRPRCLLETLRAGCLGTAEBCDLKDLAFGLURAMOTLTLDLSPNLTDAKHLCOGLRPSCKLQRLQVSPMTEGLDGTGEMASVLSASPLKELDLQNNLDVGVRLCEGLRHPACKLIRLPKSPMTEGLDGTGEMSNSSTSLKRLGSRASHVAQANLKLIDVSKIPIPIAEIAEESPEVVPVLLCVFPPASQDLHLTKPLGTDHDFWPTGVATEVVDKKNLYRVHPVAGSYRNPVTGLCFVMREAVTVEIFCVMQDQFLGEINPQHSMMVAGPLLDIKAEFGAVEAHLPHFVALQGHVDRSILPQMAHFKBEGMLKPARVELHIVLENPSFSLGVLKMIHNLARLFPVTSVVLHYHVRHPEVTFHLYLIPSCSIRKADIDLEMKFQVRIHKPPPLTPLYMGCRTVSGSGSNMILPKELELCYRSGEDQLFSEFYVHLGSGIRLOVKDKDQETLYWEALVKPGLMPATITLIPPIARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVLDKLHGQVLSQEQYERVLAEATRPSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLIMELWELKSGKKGLLPLSS"

BASE COUNT

1018 a 1208 c 1208 g 901 t

ORIGIN

Query Match

Best Local Similarity

Matches 4197; Conservative 0; Mismatches 0; Indels 225; Gaps 2;

QY

523

ATGGCTGGCGGAGCTGGGGCGCTGGCTCTTACTTGGAGTTCCTGAAGAGGAGGAG

582

Db

1

ATGGCTGGCGGAGCTGGGGCGCTGGCTCTTACTTGGAGTTCCTGAAGAGGAGGAG

60

QY

583

CTGAAGGAGTTCAGCTTCTGCTCGCCATAAAGCGCACTCCAGGAGCTTTCGGGTGAG

642

Db

61

CTGAAGGAGTTCAGCTTCTGCTCGCCATAAAGCGCACTCCAGGAGCTTTCGGGTGAG

120

QY

643

ACACCGCTCAGCCAGAGAAAGCAGTGGCAGTGGAGTGGCTCTGCTGCTGCTGAG

702

Db

121

ACACCGCTCAGCCAGAGAAAGCAGTGGCAGTGGAGTGGCTCTGCTGCTGCTGAG

180

QY

703

TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG

762

Db

181

TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG

240

QY

763

TCACTGTGGCGCCAAAGCCAGGAGGGCAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA

822

Db

241

TCACTGTGGCGCCAAAGCCAGGAGGGCAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA

300

QY

823

AGTGAACCCCACTGGGGTCTCCAGCCCAACCCCACTCCACCGCAGTGTCTAATGCCCTGG

882

Db

301

AGTGAACCCCACTGGGGTCTCCAGCCCAACCCCACTCCACCGCAGTGTCTAATGCCCTGG

360

QY

883

ATCCATGAATTCGGCGGGGTCCACCCAGGCTCAGAGAGAGGGTTTGTGACACAGCTG

942

Db

361

ATCCATGAATTCGGCGGGGTCCACCCAGGCTCAGAGAGAGGGTTTGTGACACAGCTG

420

QY

943

CCTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCTCTCCTCTTACCAAGCTCTT

1002

Db

421

CCTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCTCTCCTCTTACCAAGCTCTT

480

QY

1003

CCAAGCTCCCAAGACCATGAGTCTCCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA

1062

Db 481 CCAAGCTCCCCAGACCATGAGTCTTCAAGCCAGGAGTCAACCAAGCCGCCCATCCACA 540
 Qy 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGCAGGAG 1122
 Db 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGCAGGAG 600
 Qy 1123 GGTCTCTGGGACCAATGGGCTCTGGATGAAAGGTCAAGGAATTTACTTACACAGAAATTCAGA 1182
 Db 601 GCTCTCTGGGACCAATGGGCTCTGGATGAAAGGTCAAGGAATTTACTTACACAGAAATTCAGA 660
 Qy 1183 GAAAGAGAGAGAGAGAAATTCAGAGAAAGGAGGCGCCCAATGGGCGCGGTGGTGAAGAG 1242
 Db 661 GAAAGAGAGAGAGAGAAATTCAGAGAAAGGAGGCGCCCAATGGGCGCGGTGGTGAAGAG 720
 Qy 1243 CCCCCACAGGCGCACACAGCCTACAGCCCCCACCACCCACCCATGGGAGCCTTCTGTGAGA 1302
 Db 721 CCCCCACAGGCGCACACAGCCTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGA 780
 Qy 1303 GAGAGCCTCTGTTCCACATGGGCGCTGGAAAAATGAGGATTTTAAACCAAAAAATTCACACAG 1362
 Db 781 GAGAGCCTCTGTTCCACATGGGCGCTGGAAAAATGAGGATTTTAAACCAAAAAATTCACACAG 840
 Qy 1363 CTGCTACTTCTACAAAGACCTCAACCCAGAGCCTGCTGCTCAAGAGAGAGCTGG 1422
 Db 841 CTGCTACTTCTACAAAGACCTCAACCCAGAGCCTGCTGCTCAAGAGAGAGCTGG 900
 Qy 1423 CTTGATATGTGGAGAGAGATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 1482
 Db 901 CTTGATATGTGGAGAGAGATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 960
 Qy 1483 GGCTCTGGATACCAAGAACCTCGCATAGTCAATCTGACGGGCGCTCTGGAATTTGGGAAG 1542
 Db 961 GGCTCTGGATACCAAGAACCTCGCATAGTCAATCTGACGGGCGCTCTGGAATTTGGGAAG 1020
 Qy 1543 TCAACACTGGCCAGCGAGTGAAGAGCCTGGGGAGAGGCGAGCTGATGGGAGACCGC 1602
 Db 1021 TCAACACTGGCCAGCGAGTGAAGAGCCTGGGGAGAGGCGAGCTGATGGGAGACCGC 1080
 Qy 1603 TTCCAGCATGTCTTACTTACGTCTCAGAGAGCTGGCCAGTCCCAAGGTGGTGAAGTCTC 1662
 Db 1081 TTCCAGCATGTCTTACTTACGTCTCAGAGAGCTGGCCAGTCCCAAGGTGGTGAAGTCTC 1140
 Qy 1663 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
 Db 1141 GCTGAGCTCATCGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
 Qy 1723 AGGCCAGAGCGGTCTCTTCACTCTCGATGTGTAGATGAGCCAGGATGGGTCTTGACAG 1782
 Db 1201 AGGCCAGAGCGGTCTCTTCACTCTCGATGTGTAGATGAGCCAGGATGGGTCTTGACAG 1260
 Qy 1783 GAGCCGAGTTCTGAGTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCATGCTG 1842
 Db 1261 GAGCCGAGTTCTGAGTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCATGCTG 1320
 Qy 1843 GGCAGTTTCTGGGAAAACTACTTCTCCGAGGCACTCTCTGATCAGCGCTCGGACC 1902
 Db 1321 GGCAGTTTCTGGGAAAACTACTTCTCCGAGGCACTCTCTGATCAGCGCTCGGACC 1380
 Qy 1903 ACAGCTCTCAGAACCTCACTCTCTTTGGAGCAGGCACTGTGGGTAGAGGTCTCTGGGG 1962
 Db 1381 ACAGCTCTCAGAACCTCACTCTCTTTGGAGCAGGCACTGTGGGTAGAGGTCTCTGGGG 1440
 Qy 1963 TTCTCTGAGTCCAGAGAGAGGATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 2022
 Db 1441 TTCTCTGAGTCCAGAGAGAGGATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 1500
 Qy 2023 ATTTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTTGGCC 2082
 Db 1501 ATTTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTTGGCC 1560
 Qy 2083 TGGGTGTCTGGCTGGCTGCACTTGTGCTGATGACAGAGATGAAGCGGAGGAAATCTC 2142
 Db 1561 TGGGTGTCTGGCTGGCTGCACTTGTGCTGATGACAGAGATGAAGCGGAGGAAATCTC 1620

Qy 2143 ACACGTACTTCAAGACCAACCAACCTCTGTCTACATTTACCTTGCCAGGCTCTCCAA 2202
 Db 1621 ACACGTACTTCAAGACCAACCAACCTCTGTCTACATTTACCTTGCCAGGCTCTCCAA 1680
 Qy 2203 GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGTCTCTGGGTGTGAGGGCATCTGG 2262
 Db 1681 GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGTCTCTGGGTGTGAGGGCATCTGG 1740
 Qy 2263 CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGGAGAGCATGGTTAGATGGGGCCATC 2322
 Db 1741 CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGGAGAGCATGGTTAGATGGGGCCATC 1800
 Qy 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
 Db 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860
 Qy 2383 TTCAATCACTCTGTTTCCAAAGAGTCTTTTTCAGCAATGTCTATGTCTTGGAGATGAG 2442
 Db 1861 TTCAATCACTCTGTTTCCAAAGAGTCTTTTTCAGCAATGTCTATGTCTTGGAGATGAG 1920
 Qy 2443 AAGGGAGAGGTAAACATTTCTAAATTGCAATAGATTTTGGAAAAAGACCTAGAACATAT 2502
 Db 1921 AAGGGAGAGGTAAACATTTCTAAATTGCAATAGATTTTGGAAAAAGACCTAGAACATAT 1980
 Qy 2503 GGAATACATGGCTCTTGGGGCATCAACACAGTTTCTTATTTGGGCTGTCTTAAAGTAT 2562
 Db 1981 GGAATACATGGCTCTTGGGGCATCAACACAGTTTCTTATTTGGGCTGTCTTAAAGTAT 2040
 Qy 2563 GAGGGGAGAGAGATGGAGAACATTTTCACTCCGGCTGTCTCAGGGAGGAAACCTG 2622
 Db 2041 GAGGGGAGAGAGATGGAGAACATTTTCACTCCGGCTGTCTCAGGGAGGAAACCTG 2100
 Qy 2623 ATGAGTGGGTCCGCTCCCTGAGCTGTCTGAGCCACACACTCTCTGGAGTCCCTCCAC 2682
 Db 2101 ATGAGTGGGTCCGCTCCCTGAGCTGTCTGAGCCACACACTCTCTGGAGTCCCTCCAC 2160
 Qy 2683 TGCTTGTACGAGACTCGGAAACAAACAGTTCTCTGACACAAGTGTGAGCCCATTTTCAAGAA 2742
 Db 2161 TGCTTGTACGAGACTCGGAAACAAACAGTTCTCTGACACAAGTGTGAGCCCATTTTCAAGAA 2220
 Qy 2743 ATGGCATGTGTGTAGAAAAAGACATGAGTCTTTAGTGTGCACTTTTGTGATTTAAATTC 2802
 Db 2221 ATGGCATGTGTGTAGAAAAAGACATGAGTCTTTAGTGTGCACTTTTGTGATTTAAATTC 2280
 Qy 2803 AGCCGCAAGTGAAGAGCTTTCAGTGTATGAGGCGAGCAGCAGATCAACATGGAGC 2862
 Db 2281 AGCCGCAAGTGAAGAGCTTTCAGTGTATGAGGCGAGCAGCAGATCAACATGGAGC 2340
 Qy 2863 CCCACATGGTGTCTCTTCAGGTGGGTCCAGTCAAGATGCCTATTTGGCAGATTTCTC 2922
 Db 2341 CCCACATGGTGTCTCTTCAGGTGGGTCCAGTCAAGATGCCTATTTGGCAGATTTCTC 2400
 Qy 2923 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAGGAGCTGGAACCTAAGTGGAAAACTCGCTG 2982
 Db 2401 TTCTCCGCTCTCAAGGTCAACAGAAACCTGAGGAGCTGGAACCTAAGTGGAAAACTCGCTG 2460
 Qy 2983 AGCCTCTGCAAGTGAAGAGTCTTTGTAAAGCCCTGAGAGCCCTCGCTGCTCTCTGAG 3042
 Db 2461 AGCCTCTGCAAGTGAAGAGTCTTTGTAAAGCCCTGAGAGCCCTCGCTGCTCTCTGAG 2520
 Qy 3043 ACCCTGGGTGGCTGGCTCAGCTGAGGAGTCAAGGAGCTTGCCTTTGGG 3102
 Db 2521 ACCCTGGGTGGCTGGCTCAGCTGAGGAGTCAAGGAGCTTGCCTTTGGG 2580
 Qy 3103 CTGAGAGCAACACAGACCCCTGACCGAGCTGGAACCTTCAATGTGCTCACCGATGCT 3162
 Db 2581 CTGAGAGCAACACAGACCCCTGACCGAGCTGGAACCTTCAATGTGCTCACCGATGCT 2640
 Qy 3163 GGAGCAAAACCTTTTGGCAGAGACTGAGACCGGAGCTGCAAGCTACAGGAGCTGAG 3222
 Db 2641 GGAGCAAAACCTTTTGGCAGAGACTGAGACCGGAGCTGCAAGCTACAGGAGCTGAG 2700


```
QY 3223 CTGGTCAGCTGGGCTCAAGTCTGAGCTGCTGCGCAGGACCTGGCCTCTGTGCTTATGTGCC 3282
Db 2701 CTGGTCAGCTGGGCTCAAGTCTGAGCTGCTGCGCAGGACCTGGCCTCTGTGCTTATGTGCC 2760
QY 3283 AGCCCCAGCTGAAGAGCTAGACCTGCGCAGCAGAACCTCGATCAGCTTGGCGTGCGA 3342
Db 2761 AGCCCCAGCTGAAGAGCTAGACCTGCGCAGCAGAACCTCGATCAGCTTGGCGTGCGA 2820
QY 3343 CTGCTCTGTGAGGGGCTCAGGCACTCTGCTGCAAACTCATACGCTCTGGGCTGGACCA 3402
Db 2821 CTGCTCTGTGAGGGGCTCAGGCACTCTGCTGCAAACTCATACGCTCTGGGCTGGACCA 2867
QY 3403 ACAACTCTGAGTATGATGAGGACAGGAACTGAGGGCCCTGGAGCAGAGAGAACTCAG 3462
Db 2868 ----- 2867
QY 3463 CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTCAGGGCCTGATAGC 3522
Db 2868 -----GAAACCAAGTGTGATGACCCCTACTCAGGGCCTGATAGC 2907
QY 3523 GGAGAGATGAGTAATAGCACATCTCACTCAAGCGCGCAGAGACTCGGATCAGAGAGGGCG 3582
Db 2908 GGAGAGATGAGTAATAGCACATCTCACTCAAGCGCGCAGAGACTCGGATCAGAGAGGGCG 2967
QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCGTGAGCAAGATCTTCCCAATT 3642
Db 2968 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCGTGAGCAAGATCTTCCCAATT 3027
QY 3643 GCTGAGATGAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCT 3702
Db 3028 GCTGAGATGAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCT 3087
QY 3703 TCTCTGCTCTCAAGGGGACCTGATACGAAGCCCTTTGGGACTGACGATCACTTCTGG 3762
Db 3088 TCTCTGCTCTCAAGGGGACCTGATACGAAGCCCTTTGGGACTGACGATCACTTCTGG 3147
QY 3763 GCGCCCAAGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3148 GCGCCCAAGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3207
QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
Db 3208 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3267
QY 3883 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTTCTCTGGTGAGATCAAC 3942
Db 3268 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTTCTCTGGTGAGATCAAC 3327
QY 3943 CCACAGCAGCTGGATGGTGCGAGGGCTCTGCTGGACATCAAGGCTCAGGCTGGAGCT 4002
Db 3328 CCACAGCAGCTGGATGGTGCGAGGGCTCTGCTGGACATCAAGGCTCAGGCTGGAGCT 3387
QY 4003 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGACACATCC 4062
Db 3388 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGACACATCC 3447
QY 4063 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 4122
Db 3448 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 3507
QY 4123 GAGCTGCATCACAATGTTCTGGAACACCCAGCTTCTCCCGCTTGGGAGTCTCTCTGAAA 4182
Db 3508 GAGCTGCATCACAATGTTCTGGAACACCCAGCTTCTCCCGCTTGGGAGTCTCTCTGAAA 3567
QY 4183 ATGATCCATAATGCTCGGCTTCAATCCGCTCACTCTGTGCTGCTGTGTTTACACCGC 4242
Db 3568 ATGATCCATAATGCTCGGCTTCAATCCGCTCACTCTGTGCTGCTGTGTTTACACCGC 3627
QY 4243 GTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTACGCTTCCATTCGG 4302
Db 3628 GTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTACGCTTCCATTCGG 3687
QY 4303 ----- 4302
```

```
Db 3688 AAGGCATAGATGATCTAGAAATGAATTCAGTTTGTGCGAATCCACAAGCCACCCCG 3747
QY 4303 ----- 4302
Db 3748 CTGACCCCACTTTATATGGGCTGTCTTACACTGTGTCTGGTCTGGTTCAGGGATCTG 3807
QY 4303 -----AAGGAACTGGAGCTCTCTATCGAAAGCCCTGGAGAAAGACAGCTGTC 4350
Db 3808 GAAATACTCCCCAAGGAACTGGAGCTCTGCTATCGAAGCCCTGGAGAAAGACAGCTGTC 3867
QY 4351 TCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTCAAGTGAAGACAGAA 4410
Db 3868 TCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTCAAGTGAAGACAGAA 3927
QY 4411 GATGAGATCTGTGTGGAGGCTTGTGTAACAGAGAGATCTCATGCTTCAACTACT 4470
Db 3928 GATGAGATCTGTGTGGAGGCTTGTGTAACAGAGAGATCTCATGCTTCAACTACT 3987
QY 4471 CTGATCTCCAGCCGATAGCCCTTACCTTCACTCTGGATGCCCCGAGTTGCTGCAC 4530
Db 3988 CTGATCTCCAGCCGATAGCCCTTACCTTCACTCTGGATGCCCCGAGTTGCTGCAC 4047
QY 4531 TTTGTGGACCATGATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTTTC 4590
Db 4048 TTTGTGGACCATGATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTTTC 4107
QY 4591 GACAACTGCACTGGACAGCTGCTGAGCCAGGAGCAGTACGAGAGGTTGCTGCTGAGAAC 4650
Db 4108 GACAACTGCACTGGACAGCTGCTGAGCCAGGAGCAGTACGAGAGGTTGCTGCTGAGAAC 4167
QY 4651 ACGAGCCCGCAGCAGATCGGAAAGCTGTTTCACTTTGAGCCAGTCTCTGGACCGGAGTGC 4710
Db 4168 ACGAGCCCGCAGCAGATCGGAAAGCTGTTTCACTTTGAGCCAGTCTCTGGACCGGAGTGC 4227
QY 4711 AAAGATGGAATCTCAAGCCCTGAAGGAGACCCATCTCACTCATTTATGAACTCTG 4770
Db 4228 AAAGATGGAATCTCAAGCCCTGAAGGAGACCCATCTCACTCATTTATGAACTCTG 4287
QY 4771 GAGAGGGCGCAAAAGGAGCTCTGCGCACTCAGCAGCTGA 4812
Db 4288 GAGAGGGCGCAAAAGGAGCTCTGCGCACTCAGCAGCTGA 4329
```

```
RESULT 14
HSM800983
LOCUS HSM800983 4123 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFp586O1822 (from clone DKFp586O1822);
partial cds.
ACCESSION AL117470
VERSION AL117470.1 GI:5911938
KEYWORDS
```

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4123)

AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.

JOURNAL Direct Submission

COMMENT Submitted (15-SEP-1999) MIPS, Am Klopferstr.18a, D-82152

Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German

Genome Project.

This clone (DKFp586O1822) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available

at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
Location/Qualifiers
1..4123

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFp586O1822"
/tissue_type="uterus"
/clone_lib="586 (synonym: hutel). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1..4123
/gene="DKFp586O1822"
CDS
<1..3580
/gene="DKFp586O1822"
/note="X1AA0926, partly , alternative spliced"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAB55945.1"
/db_xref="GI:5911939"
/translation="LILQPHRPSQDPLVKRSPDYBENRHLIEIRDLFGPGLDQTQ
EPRIVLQAGAGIGKSLARQVKEAWRGQLYGRFQHVYFSCRELAKSQSVLSAEL
IGKGTATPAPRIOLISRPRLILLDGVDEPGWLQBPSSSELCHWSPQPADALLG
SILGKTLPEASFLITARTALONLIPSLQARWEVLGFSSESKRYFYRPTDERQ
A1ARFLVSKNELWALCPWPVWMLACTCLMQMKREKTLTKITLILCHLQIAQ
ALQAPLQGLRDLCSLAAGWIKTLPSPDLRKHGDLGAILISTFLMGLIQLBHPH
PLSYFHLICFEFFAAMSYLEDEKRGKHSNCIIDLKTLKLEAYGHLGFEASTRF
LIGLLSIDGERMGNIHCLRLSQGNLMOWPSLQILLPHSLHSLHCLFYERNKTFI
TOVMAHFEMGMCVETMELLVCTFCIKFSRHVKLQILEGRHRSHTWSPTMVLFRW
VPTDAYWQILFSLVTRNLKELDLGNSLSHSVKSILCTLRPRCLLETLRLAGC
GLTADCKDLALGRANQTLIELDLSFNVLDTAGAKHUCQRLQPSCKLQRLQLVSCG
LTSDDCCQDLASVSLKELLDQNNIDVDVGRLLCEGLRHPACKLIRLGLDQTL
SDMEQLRSLQKSPQLLIFSRKPSVMTPTGELDTGEMSNSTSLKRLQIGSRBA
SHVAQNLKLDVSKIPFIAEIAESSPEVPVELLCPSPASQGLDHTKPLGTDDDF
WGPTGVATVVDKKNLVRHVPVAGSVWNTGLCFVMEAVTVEIFCVMDQFLG
E1NPOHSMWAVPLDIDKAEPAVEAVHLPEFVALQGGHVDTSLQVAFHFEKGLLE
KEARVELHIVLENSPLGVLKMLHNAIRFIPVTSVLLYHVRHVEEVTFLYLI
RSPGDLQSEFYVDVQREGIARVTSVEVLDKLHGQVLSQEQYERVLAEINTRPSQMR
PSPLDAPQLLHVDFVQREGIARVTSVEVLDKLHGQVLSQEQYERVLAEINTRPSQMR
KLFLSLQSDWDRCKDGLYQALKETHPHLIMELWEGSKKGLLPLSS"
4057..4062
/polya_signal
/gene="DKFp586O1822"
4081
/polya_site
/gene="DKFp586O1822"
BASE COUNT 1002 a 1086 c 1136 g 899 t
ORIGIN

Query Match 69.9%; Score 3807.8; DB 9; Length 4123;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 3951; Conservative 0; Mismatches 2; Indels 132; Gaps 1;

QY 1365 GCTACTTCTACAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGAGAGCTGGCC 1424
Db 1 GCTACTTCTACAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGAGAGCTGGCC 60

QY 1425 TGATTATGTGGAGGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCGAG 1484
Db 61 TGATTATGTGGAGGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCGAG 120

QY 1485 CTTGGATACCCAGAAACCTTCGATAGTATCATCTGACGGGGGCTGTGGAAATTTGGGAAGTC 1544
Db 121 CTTGGATACCCAGAAACCTTCGATAGTATCATCTGACGGGGGCTGTGGAAATTTGGGAAGTC 180

QY 1545 AACACTGGCCAGGAGTCAAGGAGCTTGGGGGAGAGCCAGCTGTATGGGGAGCCGCTT 1604
Db 181 AACACTGGCCAGGAGTCAAGGAGCTTGGGGGAGAGCCAGCTGTATGGGGAGCCGCTT 240

QY 1605 CCAGCATGTCTTCTACTTCAGCTGCAGAGAGCTGGCCCGCCAGTCCAAGGTGGTGGTCTCGC 1664
Db 241 CCAGCATGTCTTCTACTTCAGCTGCAGAGAGCTGGCCCGCCAGTCCAAGGTGGTGGTCTCGC 300

QY 1665 TGAGCTCATCGGAAAGATGGAGACCCACTCGGGCTCCCATTTAGACAGATCCTGTCTAG 1724
Db 301 TGAGCTCATCGGAAAGATGGAGACCCACTCGGGCTCCCATTTAGACAGATCCTGTCTAG 360

```

```

QY 1725 GCCAGAGCGGTGCTCTTTTCATCTCGATGGTGTAGATCAGCCAGGATCGGTCTTTCAGGA 1784
Db 361 GCCAGAGCGGTGCTCTTTTCATCTCGATGGTGTAGATCAGCCAGGATGGGTCTTTCAGGA 420

QY 1785 GCCGAGTTCTGAGCTCTGTCTGCACTGAGCAGCCACAGCGCGGGATGCACTCTCTGGG 1844
Db 421 GCCGAGTTCTGAGCTCTGTCTGCACTGAGCAGCCACAGCGCGGGATGCACTCTCTGGG 480

QY 1845 CAGTTTGTCTGGGAAAACTATATCTTCCGAGGACATCTTCTCTGATCAGCGTTCGGACAC 1904
Db 481 CAGTTTGTCTGGGAAAACTATATCTTCCGAGGACATCTTCTCTGATCAGCGTTCGGACAC 540

QY 1905 AGCTCTGCAGAACCTCATTTCTCTTTTGGAGCAGCAGCTTTGGGTAGAGGTCTCTGGGTT 1964
Db 541 AGCTCTGCAGAACCTCATTTCTCTTTTGGAGCAGCAGCTTTGGGTAGAGGTCTCTGGGTT 600

QY 1965 CTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGCAAGCAAT 2024
Db 601 CTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAGCAAGCAAT 660

QY 2025 TAGAGCCCTTATAGTTGGTCAATCAAAACAAGAGCTCTGGGCCCTGTCTTGTGCCCTG 2084
Db 661 TAGAGCCCTTATAGTTGGTCAATCAAAACAAGAGCTCTGGGCCCTGTCTTGTGCCCTG 720

QY 2085 GGTGTCTCTGGCTGGCTGCACCTTGCTGATGAGCAGATGAAGCGGAAGAAAACTCAC 2144
Db 721 GGTGTCTCTGGCTGGCTGCACCTTGCTGATGAGCAGATGAAGCGGAAGAAAACTCAC 780

QY 2145 ACTGACTTCCAGAGCACCAACCCCTCTGTCTACATTTACCTTGCCTCAGGCTCTCAAGC 2204
Db 781 ACTGACTTCCAGAGCACCAACCCCTCTGTCTACATTTACCTTGCCTCAGGCTCTCAAGC 840

QY 2205 TCAGCCATTGGGACCCAGCTCAGAGACTCTGTCTCTCTGCTGCTGAGGGCATCTGGCA 2264
Db 841 TCAGCCATTGGGACCCAGCTCAGAGACTCTGTCTCTCTGCTGCTGAGGGCATCTGGCA 900

QY 2265 AAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCAT 2324
Db 901 AAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCAT 960

QY 2325 CTCACCTTCTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTCTGAGTACAGCTT 2384
Db 961 CTCACCTTCTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTCTGAGTACAGCTT 1020

QY 2385 CATTCACCTCTGTTTCCAAAGAGTCTTTCGACAGATGCTCTATGCTTTGGAGATGAGAA 2444
Db 1021 CATTCACCTCTGTTTCCAAAGAGTCTTTCGACAGATGCTCTATGCTTTGGAGATGAGAA 1080

QY 2445 GGGGAGAGGTAAACATTTCTAATTGATCATAGATTGGAAAAAGACGCTAGAGCATATGG 2504
Db 1081 GGGGAGAGGTAAACATTTCTAATTGATCATAGATTGGAAAAAGACGCTAGAGCATATGG 1140

QY 2505 AATACATCGCTGTTTGGGGCATCAACACAGTTTCTTATTTGGGCTGTTTAAGTGATGA 2564
Db 1141 AATACATCGCTGTTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGTTTAAGTGATGA 1200

QY 2565 GGGGAGAGAGATGAGGAACATCTTTTCACTGCGGCTGCTCAGGGAGGAACTCAT 2624
Db 1201 GGGGAGAGAGATGAGGAACATCTTTTCACTGCGGCTGCTCAGGGAGGAACTCAT 1260

QY 2625 GCAGTGGGTCCCGTCCCTGCGAGCTGCTGCTGAGCCACACTCTCTGGAGTCCCTCCACTG 2684
Db 1261 GCAGTGGGTCCCGTCCCTGCGAGCTGCTGCTGAGCCACACTCTCTGGAGTCCCTCCACTG 1320

QY 2685 TTGTGACGAGACTCGGAAACAAAACGTTCTGACACAAGTATGGCCCATTTTCGAAGAAT 2744
Db 1321 TTGTGACGAGACTCGGAAACAAAACGTTCTGACACAAGTATGGCCCATTTTCGAAGAAT 1380

QY 2745 GGGCATGTGTAGAAAAACAGACATGAGCTCTTAGTGTGCACTTTCTCATTTAAATTCAG 2804
Db 1381 GGGCATGTGTAGAAAAACAGACATGAGCTCTTAGTGTGCACTTTCTCATTTAAATTCAG 1440

QY 2805 CCGCCACGTGAAGAGCTTTCAGCTGATTGAGGGCAGGACGACAGATCAACATGGAGCCC 2864

```

||||| 1441 CCGCCACGTGAAGAGCTTCAGCTGATTTAGGGCAGGCAGACAGATCAACATGAGGCC 1500
QY
2865 CACCATGTGTAGTCCCTGTTTCCAGGTGGGTCCAGTCAAGATGCTATTTGGCAGATTCCTTT 2924
Db
1501 CACCATGTGTAGTCCCTGTTTCCAGGTGGGTCCAGTCAAGATGCTATTTGGCAGATTCCTTT 1560
QY
2925 CTCCTGCTCCAAAGGTGACCCAGAAAACCTGAAGAGCTGGACCTAAGTGGAAAACCTGCTGAG 2984
Db
1561 CTCCTGCTCCAAAGGTGACCCAGAAAACCTGAAGAGCTGGACCTAAGTGGAAAACCTGCTGAG 1620
QY
2985 CCACCTGTCAGTGAAGAGTCTTTGTAAGACCTGAGACGCCCTCGCTGCTCCTGGAGAC 3044
Db
1621 CCACCTGTCAGTGAAGAGTCTTTGTAAGACCTGAGACGCCCTCGCTGCTCCTGGAGAC 1680
QY
3045 CTCGGGTGGGTGGCTGTGGCTCACAGCTGAGGACTCAAGGACCTTTGCCCTTTGGGCT 3104
Db
1681 CTCGGGTGGGTGGCTGTGGCTCACAGCTGAGGACTCAAGGACCTTTGCCCTTTGGGCT 1740
QY
3105 GAGAGCCAAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAACGATGCTGG 3164
Db
1741 GAGAGCCAAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAACGATGCTGG 1800
QY
3165 AGCCAAAACCTTTGCCAGAGACTGAGACAGCGGAGCTGCAAGCTACAGCGACTGCGACT 3224
Db
1801 AGCCAAAACCTTTGCCAGAGACTGAGACAGCGGAGCTGCAAGCTACAGCGACTGCGACT 1860
QY
3225 GGTCAAGCTGTGGCTCACGCTCTGACTGTGCTGCAGGACCTGGCTCTGTCTTGTAGTGCCAG 3284
Db
1861 GGTCAAGCTGTGGCTCACGCTCTGACTGTGCTGCAGGACCTGGCTCTGTCTTGTAGTGCCAG 1920
QY
3285 CCCAGCTGGAAGGAGCTAGACTGTCAGCAGAGAACCACTGGATGACGTTGGGGTGGGACT 3344
Db
1921 CCCAGCTGGAAGGAGCTAGACTGTCAGCAGAGAACCACTGGATGACGTTGGGGTGGGACT 1980
QY
3345 GCTCTGTGAGGGGCTCAGGACTCTGCTGCGTGAACCTCATACGCTGGGGCTGGACAGAC 3404
Db
1981 GCTCTGTGAGGGGCTCAGGACTCTGCTGCGTGAACCTCATACGCTGGGGCTGGACAGAC 2040
QY
3405 AACTCTGAGTATGATGATGAGCAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCT 3464
Db
2041 AACTCTGAGTATGATGATGAGCAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCT 2100
QY
3465 GCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGGCTGGATACGGG 3524
Db
2101 GCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGGCTGGATACGGG 2160
QY
3525 AGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCGGC 3584
Db
2161 AGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCGGC 2220
QY
3585 TTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGACAGATCTTCCCAATTGC 3644
Db
2221 TTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGACAGATCTTCCCAATTGC 2280
QY
3645 TGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGGTGGCTTC 3704
Db
2281 TGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGGTGGCTTC 2340
QY
3705 TCTGCTCTCAAGGGGAACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGG 3764
Db
2341 TCTGCTCTCAAGGGGAACCTGCATACGAAGCTTTGGGGACTGACGATGACTTCTGGGG 2400
QY
3765 CCCACGGGGCTGTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCA 3824
Db
2401 CCCACGGGGCTGTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCA 2460
QY
3825 CTTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAGA 3884
Db
2461 CTTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAGA 2520
QY
3885 AGCGGTGACGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCC 3944
|||||

Db
2521 AGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACAGTTCTCTGGGTGAGATCAACCC 2580
QY
3945 ACAGCACAGCTGGATGTTGGCAGGGCCTCTGCTGGACATCAAGGTGAGCCTGGAGCTGT 4004
Db
2581 ACAGCACAGCTGGATGTTGGCAGGGCCTCTGCTGGACATCAAGGTGAGCCTGGAGCTGT 2640
QY
4005 GGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCT 4064
Db
2641 GGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGCCATGTGGACACATCCCT 2700
QY
4065 GTTCCAAATGGCCCACTTTTAAAGAGAGGGGATGCTCCTGGAGAGCCAGCCAGGGTGA 4124
Db
2701 GTTCCAAATGGCCCACTTTTAAAGAGAGGGGATGCTCCTGGAGAGCCAGCCAGGGTGA 2760
QY
4125 GCTGATCACATAGTTCCTGAAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAAAT 4184
Db
2761 GCTGATCACATAGTTCCTGAAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAAAT 2820
QY
4185 GATCATTAATGCCCTGGCTTCATTTCCCGTCACTCTCTGTGTGTTGCTTTACACCGCT 4244
Db
2821 GATCATTAATGCCCTGGCTTCATTTCCCGTCACTCTCTGTGTGTTGCTTTACACCGCT 2880
QY
4245 CCATCTCAGGAAGTCACTTCCACTCTACCTGATCCCAAGTGAAGTCTCCATTCGG-- 4302
Db
2881 CCATCTCAGGAAGTCACTTCCACTCTACCTGATCCCAAGTGAAGTCTCCATTCGGAA 2940
QY
4303 ----- 4302
Db
2941 GGCCATAGATGATCTAGAAATGAATTCAGATTGTGGGAATCCACAGCCACCCCGCT 3000
QY
4303 ----- 4302
Db
3001 GACCCCACTTTATATGGGCTGTCTGTCTGTCTGTGGTTCAAGGATGCTGGA 3060
QY
4303 -----AAGGAATGGAGCTCTGCTATCGAAGCCCTGGAGAGACAGCTGTCTC 4352
Db
3061 AATACTCCCAAGGAATGGAGCTCTGCTATCGAAGCCCTGGAGAGACAGCTGTCTC 3120
QY
4353 GGAGTCTTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAAAGA 4412
Db
3121 GGAGTCTTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAAAGA 3180
QY
4413 TGAGACTCTGTGTGGGAGGCTTGGTGAACACAGAGATCTCATGCTTCGCAACTACTCT 4472
Db
3181 TGAGACTCTGTGTGGGAGGCTTGGTGAACACAGAGATCTCATGCTTCGCAACTACTCT 3240
QY
4473 GATCCCTCAGCCCGCATAGCCGTAACCTCTGAGTCCCGGAGTCTGTCAGCTT 4532
Db
3241 GATCCCTCAGCCCGCATAGCCGTAACCTCTGAGTCCCGGAGTCTGTCAGCTT 3300
QY
4533 TGTGACACAGTATCGAGAGCAGCTGATAGCCGAGTGACATCGGTGGAGGTTGTCTTGA 4592
Db
3301 TGTGACACAGTATCGAGAGCAGCTGATAGCCGAGTGACATCGGTGGAGGTTGTCTTGA 3360
QY
4593 CAAACTGCATGACAGGCTCTGAGCCAGGAGCAGTACGAGAGGCTGCTGGCTGAGAACAC 4652
Db
3361 CAAACTGCATGACAGGCTCTGAGCCAGGAGCAGTACGAGAGGCTGCTGGCTGAGAACAC 3420
QY
4653 GAGGCCACAGCAGATGCGGAAGCTGTTGAGTTCAGCCAGTCTGCGGACCGGAAGTGA 4712
Db
3421 GAGGCCACAGCAGATGCGGAAGCTGTTGAGTTCAGCCAGTCTGCGGACCGGAAGTGA 3480
QY
4713 AGATGACTCTACCAAGCCCTGAAGGAGACCATCTCACTCACTTATGGAACCTGGGA 4772
Db
3481 AGATGACTCTACCAAGCCCTGAAGGAGACCATCTCACTCACTTATGGAACCTGGGA 3540
QY
4773 GAAGGGCAGCAAAAAGGGAACCTCTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTG 4832
Db
3541 GAAGGGCAGCAAAAAGGGAACCTCTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTG 3600
QY
4833 ACCCTTGTGCTGCTGCTTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTCTCTGCAAC 4892
Db
3601 ACCCTTGTGCTGCTGCTTTGGGTGACCTTCTTTGGGTCTCAGTTTCTTCTCTGCAAC 3660
|||||

QY 4893 AAGTTGCCATCTGTTTGCCTTCCAGCACTAAAGTAATGAATCTTTGATGATGCGCTTTC 4952
DB 3661 AAGTTGCCATCTGTTTGCCTTCCAGCACTAAAGTAATGAATCTTTGATGATGCGCTTTC 3720
QY 4953 TGGGCATTATGTCTCATCCAGGATGCCACAGGGGGCCCCAGTCCAGGTGCGCTTAACA 5012
DB 3721 TGGGCATTATGTCTCATCCAGGATGCCACAGGGGGCCCCAGTCCAGGTGCGCTTAACA 3780
QY 5013 GCATCTCAGGGAATGTCTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCA 5072
DB 3781 GCATCTCAGGGAATGTCTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCA 3840
QY 5073 TCTGTGGCCACAGCAGCAGCAAGCCTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGA 5132
DB 3841 TCTGTGGCCACAGCAGCAGCAAGCCTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGA 3900
QY 5133 ATAGGAGGACATGGAACCATTTGCTCTGGCTGTGTCTACAGGGTGAGCCCCAAATTTGG 5192
DB 3901 ATAGGAGGACATGGAACCATTTGCTCTGGCTGTGTCTACAGGGTGAGCCCCAAATTTGG 3960
QY 5193 GGTTCAGCTGGGAGGCCACGTGGATCTTGGCTTTGTACAGGAAGATCTACAGAGCAA 5252
DB 3961 GGTTCAGCTGGGAGGCCACGTGGATCTTGGCTTTGTACAGGAAGATCTACAGAGCAA 4020
QY 5253 GCCACAGAGTAAGTGGAGGAAGTTTATTCAGAAATAAAGGAGTATCAGAGCTCTTT 5312
DB 4021 GCCACAGAGTAAGTGGAGGAAGTTTATTCAGAAATAAAGGAGTATCAGAGCTCTTT 4080
QY 5313 TAGAA 5317
DB 4081 TAAAA 4085

RESULT 15
AX089769
LOCUS AX089769 4556 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 9 from Patent WO0116170.
ACCESSION AX089769
VERSION AX089769.1 GI:13443941
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE
1
AUTHORS Reed, J.C.
TITLE Card proteins involved in cell death regulation
JOURNAL Patent: WO 0116170-A 9 08-MAR-2001;
The Burnham Institute (US)
FEATURES
source
1..4556
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1..4365
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC34903.1"
/db_xref="GI:13443942"

CDS
/translation="MAGGAWRLACVLEFLKKEELKEFOLLANKAHGRSSSGETPAQ
PEKTSMEVASLYVAQYGEQRAWDALHTWEOGLRSLCAOAGCAHSPSPYSPSE
PHLGSPQSTAVLMPWHPHELPACTQGSRRVRLQPDTSGRWRHREISALLQAL
IREREREKGRPPAAVVGTPQAHTSLQPHHPWESVRESLCTWPKWNEFDNQ
KFTQLLLQHPRSDPIVKSFPDYVEENRGLHIEIDLFGPLDTPQEPRIVILQ
AAGIGKSTLAROVKAWGRGOLYGRFQHVFSRELAOSKVSLAEILGKDGATP
APIQILSEPERLLFLDGVDEPGHVLQEPSELCLHWSQPOPOPADALLGLLGLTLP
EASFLITARTALQNLILPSLEQARWVVLGFSESSKEYFYFYFDERQALRAFLVK
SNKELWALCVLPWVSNLACTLQMKRKEKLTLSKTTTLLCHLYLAQALQASPIHL
QURDLCSLAAEGIWOKTLFSPDRLKRLGDGAIISTFLKMGILQEHPIPLSYFIHL
CFQEFFAAMSVYLVDEKGRKHSNCLIDLEKTLAYGIHLFGASTTRFLGLLSDGE
BREMENI FHLRLSQRNLMQWVPSLQLLLQPHLSLSLHCLYETRNKTFITQVMAHP

NGMCVETDMELLVCTFCIKFSRHVKKQLQIEGRQHRSTWSTWVLFPRWVPVTDAYWQ
ILFVSLKVRNLTLDLSGNSLSDSAKSLCKTLRRPRCLLETTLRLACGLTAEDCKD
LAFLGRNQTITELDLSFNVLTDGAKHLQRLQPSCKLQRLQSLGSLTDCQCDL
ASVLSASLSKELDLQNNLDDVGRLLCEGRHPACKLIRGLDQTTLSDEMROELR
ALEBEKPOLLIIFSRKRSVMTPTGLDTGEMNSSTSLKROLRSGERASRAHVQANLK
LLDVSKTPIPIAEIAEESPEVVELLCVPSPOQDLHTPLGTDFFDPLGPEGNVDV
ELIDKSTNRYSWFPTAGWYLSATGLFVLRDEVTTTAFGSMQWHLALDLOHHEQW
LVGGPLFDVTAEPBEAAEHLPHFISLQGEVDVSWFLVAHFQNEGMVLEHARVEFP
YAVLESFSLMGLILLRTASCTRLSIPITSNLIVYHPHEDIKFHLVLPSDALLTK
AIDBEERPHVRLQTSPPMEPLNFGSYIVSNANLKVMPKELKLSVRSCEIOHFS
KPYAGKMEPIQLIEITERHGTLLVWDTEVPKVDLQLVAAAPPFPSPGAAFKVKNRQL
QARMGDLKGLVLDLDDQNEVLENEKELVEQETQSKNEALLSMVVEKKGLDALDLVLR
SISERDPVLVSYLRQNL"

BASE COUNT 1100 a 1223 c 1256 g 977 t
ORIGIN

Query Match 62.5%; Score 3400.4; DB 6; Length 4556;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 3555; Conservative 0; Mismatches .236; Indels 3; Gaps 1;

QY 523 ATGGCTGGCGAGCTGGGGCGCTGGCTGTCTTACTTTGGAGTTCCTGAAGAAGGAGGAG 582
DB 1 ATGGCTGGCGAGCTGGGGCGCTGGCTGTCTTACTTTGGAGTTCCTGAAGAAGGAGGAG 60
QY 583 CTGAAGAGATTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
DB 61 CTGAAGAGATTCCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCACAGAGAAAGAGTGGCATTGGAGTGGCTTCGTACTCTGTGGCTCAG 702
DB 121 ACACCCGCTCAGCCACAGAGAAAGAGTGGCATTGGAGTGGCTTCGTACTCTGTGGCTCAG 180
QY 703 TATGGGAGCAGCGGGCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGCTGAGG 762
DB 181 TATGGGAGCAGCGGGCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGCTGAGG 240
QY 763 TCACCTGTGGCCCAAGCCAGAGAGGGGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA 822
DB 241 TCACCTGTGGCCCAAGCCAGAGAGGGGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGGCCAACCCACTCCACCGCAGTGTGTAATGCCCTGG 882
DB 301 AGTGAACCCCACTGGGGTCTCCAGGCCAACCCACTCCACCGCAGTGTGTAATGCCCTGG 360
QY 883 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGAAGGTTTGAAGACAGCTG 942
DB 361 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGAAGGTTTGAAGACAGCTG 420
QY 943 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTT 1002
DB 421 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCCAGCCCAATCCACA 1062
DB 481 CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCCAGCCCAATCCACA 540
QY 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCACCCAGAGAGGAGGAG 1122
DB 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCACCCAGAGAGGAGGAG 600
QY 1123 GCTCTGGGACCCCAATGCGCTCTGGATGAAACGTGAGAAATTTACTACACAGAAATCAGA 1182
DB 601 GCTCTGGGACCCCAATGCGCTCTGGATGAAACGTGAGAAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGGACAGCGGTGTGAGAACG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGGACAGCGGTGTGAGAACG 720
QY 1243 CCCCCACAGGCGCACACAGGCTACAGCCCAACACACACCCATGGAGCCTTCTCTGAGA 1302
DB 721 CCCCCACAGGCGCACACAGGCTACAGGCCCCACACACCCATGGAGCCTTCTCTGAGA 780
QY 1303 GAGAGCCTCTGTTCCATCATGCGCCCTGGGAAATAGGATTTTAAACCAAAATTCACACAG 1362

Db	781	 GAGAGCCTCTGTTTCCACATGGCCCTGGAAAAATGAGATTTTAAACCAAAAAATTCACACAG	840
Qy	1363	CTGCTACTTCTCAAAAGACCTCACCCCAGAAAGCCAAAGATCCCTCTGGTCAAGAGAAGCTGG	1422
Db	841	CTGCTACTTCTCAAAAGACCTCACCCCAGAAAGCCAAAGATCCCTCTGGTCAAGAGAAGCTGG	900
Qy	1423	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAAGAGACTTTATTGGCCCA	1482
Db	901	CCTGATTATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAAGAGACTTTATTGGCCCA	960
Qy	1483	GGCTGATATCCNAGAAACCTCGCATAGTCATACTGCAGGGGCTGCTGAAATTGGGAAG	1542
Db	961	GGCTGATATCCNAGAAACCTCGCATAGTCATACTGCAGGGGCTGCTGAAATTGGGAAG	1020
Qy	1543	TCAACACTGGCCAGGCGAGGTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCCG	1602
Db	1021	TCAACACTGGCCAGGCGAGGTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCCG	1080
Qy	1603	TTCCAGCATGTCTTCTACTTTTCAGCTGCAGAGAGCTGCGCCAGTCCAAAGTGGTGAAGTCT	1722
Db	1081	TTCCAGCATGTCTTCTACTTTTCAGCTGCAGAGAGCTGCGCCAGTCCAAAGTGGTGAAGTCT	1140
Qy	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCCTGTCT	1722
Db	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTAAGACAGATCCTGTCT	1200
Qy	1723	AGGCCAGAGCGGCTGCTTCATCCTCGATGGTGTAGATGAGCCAGGATCGGTCTTGCGAG	1782
Db	1201	AGGCCAGAGCGGCTGCTTCATCCTCGATGGTGTAGATGAGCCAGGATCGGTCTTGCGAG	1260
Qy	1783	GAGCCGAGTTCTGAGCTCTGCTGCACCTGGAGCCAGCCACAGCCGCGGATGCACCTGCTG	1842
Db	1261	GAGCCGAGTTCTGAGCTCTGCTGCACCTGGAGCCAGCCACAGCCGCGGATGCACCTGCTG	1320
Qy	1843	GGCAGTTTGTGGGAAAACTATATCTTCCGAGGCATCTTCTCTGATCACGGTCTGGACC	1902
Db	1321	GGCAGTTTGTGGGAAAACTATATCTTCCGAGGCATCTTCTCTGATCACGGTCTGGACC	1380
Qy	1903	ACAGCTCTGCAGAACCTCATCTCTTCTTGAGCAGCAGCTGTTGGGTAGAGGTCTCTGGG	1962
Db	1381	ACAGCTCTGCAGAACCTCATCTCTTCTTGAGCAGCAGCTGTTGGGTAGAGGTCTCTGGG	1440
Qy	1963	TTCTCTGAGTCCAGCAGGAAGAAATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	2022
Db	1441	TTCTCTGAGTCCAGCAGGAAGAAATTTCTACAGATATTTTCAAGATGAAGGCAAGCA	1500
Qy	2023	ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
Db	1501	ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
Qy	2083	TGGGTGTCTGGCTGGCCCTGCATCTGCTGATGCAGCAGATGAAGCCGAGGAAAAATC	2142
Db	1561	TGGGTGTCTGGCTGGCCCTGCATCTGCTGATGCAGCAGATGAAGCCGAGGAAAAATC	1620
Qy	2143	ACACTGACTTCCAAGACCAACACAAACCTCTGTCTACATTACCTTGCCAGGCTCTCAA	2202
Db	1621	ACACTGACTTCCAAGACCAACACAAACCTCTGTCTCTATTTACCTTGCCAGGCTCTCAA	1680
Qy	2203	GCTCAGCCATTGGGACCCCAAGTCTGCTCTGCTGATGAGGATGAAGGCAATCTGG	2262
Db	1681	GCTCAGCCATTGGGACCCCAAGTCTGCTCTGCTGATGAGGATGAAGGCAATCTGG	1740
Qy	2263	CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGTGTAGTGGGCCATC	2322
Db	1741	CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGTGTAGTGGGCCATC	1800
Qy	2323	ATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGAGACCCCAATCCCTCTGAGCTACAGC	2382
Db	1801	ATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGAGACCCCAATCCCTCTGAGCTACAGC	1860
Qy	2383	TTCAATCAACCTCTGTTTCCAAAGATTTCTTTGAGCAATGTCTTATGTCTTGGAGGATGAG	2442

1861	Db	TTCAATTCACTCTGTTTCCAAGAGTTCCTTTGCAGCAATGTCCTATGCTCTGGAGGATGAG	1921
2443	Qy	AAGGGGAGAGGTAAACATCTTAATTGCATCATAGATTTGGAAAGACGCTAGAGCATAT	2502
1921	Db	AAGGGGAGAGGTAAACATCTTAATTGCATCATAGATTTGGAAAGACGCTAGAGCATAT	1980
2503	Qy	GGAATATATGCGCTGTTTGGGGCATCAAACACACGTTTCTATATGGGCCCTGTTAAGTGAT	2562
1981	Db	GGAATATATGCGCTGTTTGGGGCATCAAACACACGTTTCTATATGGGCCCTGTTAAGTGAT	2040
2563	Qy	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG	2622
2041	Db	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTG	2100
2623	Qy	ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGAGTCCCTCCAC	2682
2101	Db	ATGCAGTGGGTCCCGTCCCTGCAGCTGCTGCTGCAGCCACACTCTCTGAGTCCCTCCAC	2160
2683	Qy	TGCTTTGACGAGACTCGGAACAAACGTTTCTTGACACAAAGTGATGGGCCATTTTGAAGAA	2742
2161	Db	TGCTTTGACGAGACTCGGAACAAACGTTTCTTGACACAAAGTGATGGGCCATTTTGAAGAA	2220
2743	Qy	ATGGGCATGTGTGAGAAAACAGACATCGAGCTCTTAGTGTGCATTTCTGCATTAATAATTC	2802
2221	Db	ATGGGCATGTGTGAGAAAACAGACATCGAGCTCTTAGTGTGCATTTCTGCATTAATAATTC	2280
2803	Qy	AGCCGCCACCTGAAAGAGCTTCAGCTGATTTGAGGGCAGGCAGCACAGATCAACATGGAGC	2862
2281	Db	AGCCGCCACCTGAAAGAGCTTCAGCTGATTTGAGGGCAGGCAGCACAGATCAACATGGAGC	2340
2863	Qy	CCCACCATGTGATGCTTTCAGGTGGGTCCCAGTGCACAGATGCCTATTTGGCAGATTTCTC	2922
2341	Db	CCCACCATGTGATGCTTTCAGGTGGGTCCCAGTGCACAGATGCCTATTTGGCAGATTTCTC	2400
2923	Qy	TTCTCCGTCTCAAGGTCAACGAGAACCTGAGAGAGCTGCACCTAAAGTGGAAACCTCGCTG	2982
2401	Db	TTCTCCGTCTCAAGGTCAACGAGAACCTGAGAGAGCTGCACCTAAAGTGGAAACCTCGCTG	2460
2983	Qy	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCTCGAGACCCCTCGCTGCTCCTCGGAG	3042
2461	Db	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCTCGAGACCCCTCGCTGCTCCTCGGAG	2520
3043	Qy	ACCTCGGGTTGGCTGTGGCTCACAGCTGAGGACTGCAAGGACCTTGTCCTTTGGG	3102
2521	Db	ACCTCGGGTTGGCTGTGGCTCACAGCTGAGGACTGCAAGGACCTTGTCCTTTGGG	2580
3103	Qy	CTGAGAGCCAACAGACCTTGACCCGAGCTGGACCTGTAGCTTCAATGTGCTCACGGATGCT	3162
2581	Db	CTGAGAGCCAACAGACCTTGACCCGAGCTGGACCTGTAGCTTCAATGTGCTCACGGATGCT	2640
3163	Qy	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	3222
2641	Db	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	2700
3223	Qy	CTGCTCAGCTGTGGCCCTCAGCTCTGACTGTGCTGCCAGGACCTTGCTTGTAGTGCC	3282
2701	Db	CTGCTCAGCTGTGGCCCTCAGCTCTGACTGTGCTGCCAGGACCTTGCTTGTAGTGCC	2760
3283	Qy	AGCCCCAGCTGGAAGAGACTAGACCTGCAGCAGAAACAACTGTGATGAGCTTTGGCTGGGA	3342
2761	Db	AGCCCCAGCTGGAAGAGACTAGACCTGCAGCAGAAACAACTGTGATGAGCTTTGGCTGGGA	2820
3343	Qy	CTGCTCTGTAGGGGCTCAGGCATCTCTGCCCTGCAAACTCATACCCCTGGGGCTGGACCG	3402
2821	Db	CTGCTCTGTAGGGGCTCAGGCATCTCTGCCCTGCAAACTCATACCCCTGGGGCTGGACCG	2880
3403	Qy	ACAACTCTGAGTGATGAGATGAGCGCAGGAACTGAGGGGCCCTGGAGCAGGAGAAACCTCAG	3462
2881	Db	ACAACTCTGAGTGATGAGATGAGCGCAGGAACTGAGGGGCCCTGGAGCAGGAGAAACCTCAG	2940
3463	Qy	CTGCTCATCTTCAGCAGACGAAACCAAGTGTGATGATACCCCTACTGAGGGGCTGGATACG	3522
2941	Db	CTGCTCATCTTCAGCAGACGAAACCAAGTGTGATGATACCCCTACTGAGGGGCTGGATACG	3000

QY 3523 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
Db |||||
QY 3001 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3060
Db |||||
QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCGTAGCAAGATCTTCCCAATT 3642
Db |||||
QY 3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCGTAGCAAGATCTTCCCAATT 3120
Db |||||
QY 3643 GCTGAGATTGCCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAAGCTCTTGTGCGTGCCT 3702
Db |||||
QY 3121 GCTGAGATTGCCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAAGCTCTTGTGCGTGCCT 3180
Db |||||
QY 3703 TCTCTGCTCTCTCAAGGGGACCTGTCATACGAGCCCTTTGGGACTGACGATGACTTCTGG 3762
Db |||||
QY 3181 TCTCTGCTCTCTCAAGGGGACCTGTCATACGAGCCCTTTGGGACTGACGATGACTTCTGG 3240
Db |||||
QY 3763 GGCCCAACGGGGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3822
Db |||||
QY 3241 GGCCCTGAAGGAAATGTGGATGTTGAGTTGATTGATAAGAGCACAAACAGATACAGCGTT 3300
Db |||||
QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
Db |||||
QY 3301 TGGTTCCTCCCACTGCTGGCTGGTATCTGTGTGAGCCACAGGGCTCGGCTTCTGGTAAAG 3360
Db |||||
QY 3883 GAAGCGGTGACCGTTGAGATTGAAATTTCTGTGTGGGACCAAGTTCCTGG---GTGAGATC 3939
Db |||||
QY 3361 GATGAGTCAAGTGAAGATGCGTTTGGTTCTTGGAGTCAGCACCTGGCCCTGGACCTG 3420
Db |||||
QY 3940 AACCACACACAGCTGAGTGGTGGGAGGGCCCTCTGCTGGACATCAAGGCTGAGCCTGGA 3999
Db |||||
QY 3421 CAGCACCATGAACAGTGGCTGGTGGGCGGCCCTTGTGATGTCACTGCAGAGCCAGAG 3480
Db |||||
QY 4000 GCTGTGGAAGCTGTGACCTCCCTCACTTTGTGGTCTTCAAGGGGGCCATGTGACACA 4059
Db |||||
QY 3481 GAGGCTGTCCCGGAAATCCACCTCCCCCACTTCATCTCCCTCCAGGTGAGGTGGACGTC 3540
Db |||||
QY 4060 TCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTGGAGAAAGCCAGCCAGG 4119
Db |||||
QY 3541 TCTGTGTTTCTCGTTGCCCATTTAAGATGAAGGATGATGCTGGAGCATTCAGCCCGG 3600
Db |||||
QY 4120 GTGAGCTGCATCACAATGTTCTGGAAAACCCAGCTTCTCCCCCTTGGAGTCTCTCTG 4179
Db |||||
QY 3601 GTGAGCCTTTCTATGCTGTCTGTGAAAGCCCGAGCTTCTCTGTATGGGCATCCTGCTG 3660
Db |||||
QY 4180 AAAATGATCCATNAATGCCCTGGCTTCATTCCCGTCACTCTGTGTGTGCTTTTACAC 4239
Db |||||
QY 3661 CGGATCGCCAGTGGGACTCGCCTCTCCATCCCATCACTTCCAAACACATTTGATCTATTAT 3720
Db |||||
QY 4240 CGCGTCCATCCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCCATT 4299
Db |||||
QY 3721 CACCCCAACCCGAGATATTAGTTCCACTTGACTTGCTGCTGCTGCTGCTGCTGCTA 3780
Db |||||
QY 4300 CGGAAGGAAGTGA 4313
Db |||||
QY 3781 ACAAGGGCGATAGA 3794
Db |||||